Practicable methodologies for delivering comprehensive spatial soils information

Brendan Philip Malone

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

2012

Faculty of Agriculture and Environment

The University of Sydney

New South Wales

Australia

Certificate of Originality

This thesis is submitted to the University of Sydney in fulfilment of the requirements for the Doctor of Philosophy.

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Signature: Brendan Malone

Date: 17/09/12

Thesis Summary

This thesis is concerned with practicable methodologies for delivering comprehensive spatial soil information to end-users. There is a need for relevant spatial soil information to complement objective decision-making for addressing current problems associated with soil degradation; for modelling, monitoring and measurement of particular soil services; and for the general management of soil resources. These are real-world situations, which operate at spatial scales ranging from field to global scales. As such, comprehensive spatial soil information is tailored to meet the spatial scale specifications of the end user, and is of a nature that fully characterises the whole-soil profile with associated prediction uncertainties, and where possible, both the predictions *and* uncertainties have been independently validated. 'Practicable' is an idealistic pursuit but nonetheless necessary because of a need to equip land-holders, private-sector and non-governmental stakeholders and, governmental departments including soil mapping agencies with the necessary tools to ensure wide application of the methodologies to match the demand for relevant spatial soil information. Practicable methodologies are general and computationally efficient; can be applied to a wide range of soil attributes; can handle variable qualities of data; and are effective when working with very large datasets.

In this thesis, delivering comprehensive spatial soil information relies on coupling legacy soil information (principally site observations made in the field) with Digital Soil Mapping (DSM) which comprises quantitative, state-of-the-art technologies for soil mapping. After the General Introduction, a review of the literature is given in Chapter 1 which describes the research context of the thesis. The review describes soil mapping first from a historical perspective and rudimentary efforts of mapping soils and then tracks the succession of advances that have been made towards the realisation of populated, digital spatial soil information databases where measures of prediction certainties are also expressed. From the findings of the review, in order to deliver comprehensive spatial soil information to end-users, new research was required to investigate: 1) a general method for digital soil mapping the whole-profile (effectively pseudo-3D) distribution of soil properties; 2) a general method for quantifying the total prediction uncertainties of the digital soil maps that describe the whole-profile distribution of soil properties; 3) a method for validating the whole-profile predictions of soil properties *and* the quantifications of their

uncertainties; 4) a systematic framework for scale manipulations or upscaling and downscaling techniques for digital soil mapping as a means of generating soil information products tailored to the needs of soil information users. Chapters 2 to 6 set about investigating how we might go about doing these with a succession of practicable methodologies.

Chapter 2 addressed the need for whole-profile mapping of soil property distribution. Equal-area spline depth functions coupled with DSM facilitated continuous mapping the lateral and vertical distribution of soil properties. The spline function is a useful tool for deriving the continuous variation of soil properties from soil profile and core observations and is also suitable to use for a number of different soil properties. Generally, mapping the continuous depth function of soil properties reveals that the accuracy of the models is highest at the soil surface but progressively decreases with increasing soil depth.

Chapter 3 complements the investigations made in Chapter 2 where an empirical method of quantifying prediction uncertainties from DSM was devised. This method was applied for quantifying the uncertainties of whole-profile digital soil maps. Prediction uncertainty with the devised empirical method is expressed as a prediction interval of the underlying model errors. The method is practicable in the sense that it accounts for all sources of uncertainty and is computationally efficient. Furthermore the method is amenable in situations where complex spatial soil prediction functions such as regression kriging approaches are used.

Proper evaluation of digital soil maps requires testing the predictions and the quantification of the prediction uncertainties. Chapter 4 devised two new criteria in which to properly evaluate digital soil maps when additional soil samples collected by probability sampling are used for validation. The first criterion addresses the accuracy of the predictions in the presence of uncertainties and is the spatial average of the statistical expectation of the Mean Square Error of a simulated random value (MSES). The second criterion addresses the quality of the uncertainties which is estimated as the total proportion of the study area where the (1-α)-prediction interval (PI) covers the true value (APCP). Ideally these criteria will be coupled with conventional measures of map quality so that objective decisions can be made about the reliability and subsequent suitability of a map for a given purpose. It was revealed in Chapter 4, that the quantifications of uncertainty are susceptible to bias as a result of using legacy soil data to construct spatial soil prediction functions. As a

consequence, in addition to an increasing uncertainty with soil depth, there is increasing misspecification of the prediction uncertainties.

Chapter 2, 3, and 4 thus represent a framework for delivering whole-soil profile predictions of soil properties and their uncertainties, where both have been assessed or validated across mapping domains at a range of spatial scales for addressing field, farm, regional, catchment, national, continental or global soil-related problems. The direction of Chapters 5 and 6 however addresses issues specifically related to tailoring spatial soil information to the scale specifications of the end-user through the use of scale manipulations on existing digital soil maps. What is proposed in Chapter 5 is a scaling framework that takes into account the scaling triplet of digital soil maps—extent, resolution, and support—and recommends pedometric methodologies for scale manipulation based on the scale entities of the source and destination maps. Upscaling and downscaling are descriptors for moving up to coarser or down to finer scales respectively but may be too general for DSM. Subsequently *Fine-gridding* and *coarse-gridding* are operations where the grid spacing changes but support remains unchanged. *Deconvolution* and *convolution* are operations where the support always changes, which may or may not involve changing the grid spacing. While *disseveration* and *conflation* operations occur when the support and grid size are equal and both are then changed equally and simultaneously.

There is an increasing richness of data sources describing the physical distribution of the Earth's resources with improved qualities and resolutions. To take advantage of this, Chapter 6 devises a novel procedure for downscaling, involving *disseveration.* The method attempts to maintain the mass balance of the fine scaled predictions with the available coarse scaled information, through an iterative algorithm which attempts to reconstruct the variation of a property at a prescribed fine scale through an empirical function using environmental or covariate information. One of the advantages associated with the devised method is that soil property uncertainties at the coarse scale can be incorporated into the downscaling algorithm.

Finally Chapter 7 presents a synthesis of the investigations made in Chapters 2 to 6 and summarises the pertinent findings. Directly from the investigations carried out during this project there are opportunities for further work; both in terms of addressing shortcomings that were highlighted but not investigated in the thesis, and

more generally for advancing digital soil mapping to an operational status and beyond.

Acknowledgements

To be at the stage of writing this section of the thesis and being in a reflective frame of mind, I get a sense of fulfilment at having achieved something I never knew I had the capacity to do. What has transpired; this thesis being a physical reminder, is something that I am and will probably always be truly proud of. Mind you, I could never have done this alone!

I am thankful most of all to my wife Louise. Can you believe it? Its finito. Without Louise, none of this could have been even imagined let alone contemplated. Louise is my rock, solid as any known foundation. I am lucky.

There is an old proverb that says: *Lean against a good tree, and it will shelter thee.* Alex McBratney my supervisor is a source of inspiration, possessing a mind that appears to have no limits and continually surprises me. It has been an absolute privilege to have worked in Alex's research group. It was only a chance discussion back in 2008 when the seed of this thesis was planted. Since then under Alex's tutelage, he has stretched me intellectually and has pushed me beyond what I thought were the limits of my capacities. Thanks for the laughs along the way Alex.

My co-supervisor and Alex's protégé, Budiman Minasny is especially deserving of my gratitude. I am at a loss to explain the expanse of knowledge that Budi possesses, only to say that I am a little envious. It has been an immense relief to have a sounding-board in Budi and for him to have had the patience in guiding me through the treacherous territory of spatial statistics.

To the other teaching and faculty members of the Sydney University Soil Science group- Brett Whelan, Stephen Cattle, Damien Field, James Taylor, Tom Bishop, Inakwu Odeh and Balwant Singh I express much gratitude. Without singling out anyone, you have all instilled in me a newfound capacity for enquiry and a passion of soils which dates back to my undergraduate days. The passion you all demonstrate in your areas of expertise is infectious which has ensured a stimulating environment to work in.

Coming to work each day just would not be any fun without having my friends and colleagues around. In order of proximity to my desk, thank you to Kanika Singh, Ichsani Wheeler, Rob Pallasser, Derek Yates, Konrad Muller, Michael Nelson, Uta Stockmann, Lori Watson, Adrienne Ryan, Sam Player, Irshad Bibi, Nabel Niazi and Senani Karunaratne. All of you have coloured my experiences these last few years. I have many fond memories to keep. The same goes for Grant Tranter (who showed me the ropes in my first year), Sun Wei (C++ programmer extraordinaire), Nathan Odgers, Pierre Roudier, Jason Lessels, Ed Milne, and Marcelo Stabile who have been around and shared the life at certain times these last three years.

I appreciate so much the tireless efforts of Michael Short who in my 2nd year assisted with the fieldwork in the Hunter Valley. We had plenty of laughs along the way, and by the end of our stay up there we were both adept at pulling trucks (and tractors!) out of bogs and resistant to mosquito bites. Good times Shorty, thanks heaps. I would also like to extend a very hearty thanks to the Hunter Valley landholders who allowed us onto their properties to take some of their soils. It was a stimulating experience to get out there and talk to so many people, a few of whom were completely baffled as to why someone wanted to sample their soil.

Much appreciation is extended to the Jaap de Gruijter (Wageningen) who visited our research group in 2009. The conception of the validation project occurred on a road trip back from the Hunter Valley with Alex, Jaap and Ichsani. Alex and Jaap's conversation was difficult to follow at times, but what was clearly apparent was Jaap's prodigious knowledge of sampling design and statistics. I have appreciated Jaap's wise sentiments throughout my PhD. Furthermore, I would also like to express my appreciation to another Dutchman, Dick Brus (Wageningen). Dick's timely visit in 2010 to us coincided with the data analysis work I was doing from the results of my fieldwork. Dick's expertise was invaluable and I learnt alot during his visit.

I express my gratitude to Geoff Laslett (CSIRO) whom has now sadly passed. I never met Geoff, but by all accounts he was a meticulous scientist and statistician. The work he did on formulating the mass-preserving spline depth function in 1999 with Tom Bishop was central to my PhD. I am fortunate to have co-authored the spline mapping paper with Geoff in 2009.

Getting research papers into journals has been one of the most difficult yet rewarding aspects of my PhD. It is such a valuable exercise to have your research validated by a wider audience. I have the utmost respect for the peer review process and subsequently I extend my appreciation to those anonymous reviewers who provided constructive evaluations to the research papers that have come out of this PhD.

One of the really enjoyable aspects of my PhD has been the opportunity to attend meetings, workshops and conferences. These gatherings have been great opportunities to meet interesting, inspiring and influential people in the world of soil science and digital soil mapping. My international experiences have included attending the Geomorphometry conference in Zurich, Switzerland. Tomislav Hengl's (Wageningen) R and geostatistics workshop opened my mind up to the powerful world of R and statistical computing. Thanks Tom for introducing me to the 'Happy Triangle' of Open Source software. During my European experience in 2009 I travelled to the JRC European Commission in Ispra, Italy to visit Florence Carré and Luca Montanarella whom were both very kind hosts. The 10am aperitif was truly memorable. I then visited Phillipe Lagacherie (INRA) and James Taylor in Montpellier, France whom again were exceptional hosts. Thanks James for being my tour guide and immersing me into the French culture. Montpellier is beautiful and a must to visit again sometime. In 2010 I attended the 4th Global Workshop on Digital Soil Mapping, Rome. Rome was a beautiful backdrop to a wonderful conference where I met so many people to mention. The post-conference tour around Tuscany was particularly remarkable. The DSM workshop later in 2010 at Ispra, Italy was a trip where I felt really immersed into Italian culture. Lago Maggiore in September is sensational and was a beautiful backdrop to many dinners over pizza and wine with the workshop attendees. A special thanks to Budi, Tom Hengl (ISRIC), Phillipe Lagacherie (INRA), David Rossiter (ITC), Endre Dobos (University of Miskolc, Hungary), Kabindra Adhikari (Denmark), Bob MacMillan (ISRIC), Fred Young (USDA), Amy Saunders (USDA) and Yusuf Yigini (JRC) for the wonderful times and great conversation; I learnt so much.

Also in 2010 I attended the 19th World Congress of Soil Science in Brisbane. This was a very big gathering of soil scientists and a great opportunity to meet and re-acquaint myself with many of the

delegates. I had many great evenings with my Sydney University friends around the Brisbane CBD and thanks to all those that made this trip such a memorable experience.

In 2011 I was part of organising a DSM workshop at Sydney University for delegates from various soil mapping agencies around Australia. It was fantastic to be involved in this and it presented me with an opportunity to teach and coordinate exercises. Without mentioning names, thanks to all the attendees for being such a receptive audience at this workshop. However, I would especially like to acknowledge again Budi and Ichsani for their help. Thanks also to Peter Wilson (CSIRO) and David Jacquier (CSIRO) whom organised this workshop and made it such an enjoyable experience.

I have been fortunate to have met many other prominent people in Australian soil science including Neil McKenzie (CSIRO), Mike Grundy (CSIRO) and Noel Schoknect (DAFWA) whom have directly or indirectly been involved in my project. Thanks also to Ted Griffin (DFAWA) and Karen Holmes (UWA) who have had a direct involvement in my project with some soil mapping work in Dumbleyung, WA. Similarly I would like to express gratitude to Bill Verboom (DFAWA) who drove Alex and myself around Bert Facey country in Western Australia. Bill's knowledge of the soils in this area is exceptional and I got a lot out of this brief fieldtrip. Another person who has been really helpful in my project is Darren Kidd (DPIDWE). Darren has really taken up the Digital Soil Mapping torch in Tasmania and its been great to work with such an enthusiastic guy.

My day to day experiences at Sydney University have been coloured by various teaching and tutorial roles which have extended me as a person and are too many to mention. However I would particularly like to acknowledge Alex's 4th year honours students Sophie Gulliver (2009) and Phil Hughes (2011). I worked closely with these guys on their projects and had a great and fulfilling time doing so. I think with these projects everyone learned something new.

The last of the official acknowledgments goes to the Vice Chancellor of Sydney University Michael Spence whom personally offered me the PhD scholarship (VCRS) to supplement my research here. I feel privileged to have been a recipient of a VCRS and financially it was very welcomed. I also express much gratitude to the Faculty of Agriculture, Food and Natural Resources (FAFNR), who have provided me with a comfortable and organised environment to work in. Although I have had three different offices in three years, the admin staff at FAFNR have been truly exceptional. I also thank FAFNR for the financial assistance they have provided over the three years of my study that partly covered the expenses of my trips overseas. Thankyou.

On a personal note I extend much thanks and love to my family who have supported me throughout my research. Specifically, Clair, Matthew, Simon, Damien, Mary and Gordon. You have all been so special to me in your own individual ways. All this would have been impossible without you guys.

Lastly, to my daughters Juliette and Mara, my love for you both is eternal. Juliette was born at the start of my PhD and Mara was born towards the end. Nearly a day has not gone by where I have not laughed myself silly. I am so lucky to have three adorable girls in my life!

For

Louise, Juliette, and Mara.

You know the mountain won't go falling if you're still

willing to climb

but when the mountain goes falling true riches you will find

Summas Bliss, WU LYF (2011)

Chapters of this thesis have been submitted and/or published in scientific journals:

Chapter 2-

Malone, B.P., McBratney, A.B., Minasny, B. and Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. Geoderma, 154: 138-152.

Chapter 3-

Malone, B.P., McBratney, A.B. and Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. Geoderma, 160: 614-626.

Chapter 4-

Malone, B.P., de Gruijter, J.J., McBratney, A.B., Minasny, B. and Brus, D.J., 2011. Using additional criteria for measuring the quality of predictions and their uncertainties in a digital soil mapping framework. Soil Science Society of America Journal, 75: 1032-1043.

Chapter 5-

Malone, B.P., McBratney, A.B., and Minasny, B., 2012. Some methods regarding manipulations of scale for digital soil mapping. In: B. Minasny, B.P. Malone, A.B. McBratney (Eds.), Digital Soil Assessments and Beyond. CRC Press, The Netherlands, pp. 135-138.

Chapter 6-

Malone, B.P., McBratney, A.B., Minasny, B. and Wheeler, I., 2011. A general method for downscaling earth resource information. Computers and Geosciences, 41: 119-125.

Table of Contents

Thesis Summary V

Acknowledgments IX

Table of Contents XIX

List of Figures XXIV

List of Tables XVIII

General introduction 1

Chapter 1. Context of research: a review of the literature 9

1.1. Introduction 13

1.2. Early mapping of soils 14

1.3. Mapping soil spatial variation 15

1.3.1. Conventional soil mapping 16

1.3.2. Drawbacks of conventional soil mapping 17

1.3.3. Continuous and quantitative representations of soil Variability 17

1.3.3.1. Fuzzy sets 18

1.3.3.2. Geospatial models 18

1.3.3.3. Geostatistics 19

1.3.3.4. Multivariate geospatial models 21

1.3.3.5. Digital soil mapping 22

1.4. Whole-profile digital soil mapping 26

1.5. Uncertainty estimation of digital soil maps 28

1.6. Assessing the quality of digital soil maps with validation 30

1.7. Serving the end-user: scale and manipulations thereof for digital soil mapping 31

1.8. Thesis outline 34

1.9. References 36

Chapter 2. Mapping continuous depth functions of soil carbon storage and available water capacity **43**

2.1. Introduction **48**

2.2. Materials and methods **50**

2.2.1. Study area **50**

2.2.2. Environmental data **51**

2.2.3. The equal area smoothing spline **52**

2.2.4. Mapping the smoothing spline depth function **53**

2.2.5. Model validation **56**

2.2.6. Implementation of methods **57**

2.3. Results **57**

2.3.1. Fitting equal area splines to the dataset **57**

2.3.2. Stepwise regression of environmental factors **57**

2.3.3. Neural network for prediction of carbon and AWC from environmental factors **58**

2.3.4. Model validation with the 80 withheld data points **58**

2.3.5. Mapping carbon storage and available water capacity **65**

2.3.6. Scenario-based queries of the generated soil geo-database **69**

2.4. Discussion **72**

2.5. Conclusions **74**

2.6. References **75**

Appendix 2.1. Derivation of the equal-area spline function **78**

Chapter 3. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. **83**

3.1. Introduction **88**

3.2. Theory and scope of work **90**

3.2.1. The prediction interval as a measure of uncertainty **90**

3.2.2. Validating the prediction interval **92**

3.2.3. Fuzzy clustering **93**

3.2.4. Adaptation of the UNEEC approach for digital soil mapping of continuous depth functions **95**

3.2.5. Procedure **95**

3.3. Materials and methods **98**

3.3.1. The data **98**

3.3.2. The prediction model **99**

3.3.2.1. Prediction model calibration **99**

3.3.2.2. Prediction model validation **99**

3.3.3. The empirical uncertainty model **100**

3.3.3.1. Uncertainty model calibration: fuzzy clustering and formulation of cluster prediction intervals **100**

3.3.3.2. Uncertainty model validation **101**

3.3.4. Mapping of predictions and their uncertainties **102**

3.3.5. Implementation of methods **102**

3.4. Results and discussion **102**

3.4.1. The prediction model **102**

3.4.1.1. Prediction model calibration **102**

3.4.1.2. Prediction model validation **102**

3.4.2. The empirical uncertainty model **105**

3.4.2.1. Uncertainty model calibration: fuzzy clustering and formulation of cluster prediction intervals **105**

3.4.2.2. Uncertainty model validation **108**

3.4.3. Mapping predictions and their uncertainties **110**

3.5. Conclusions **118**

3.6. References **120**

Chapter 4. Criteria and sampling for simultaneously measuring the quality of predictions *and* their uncertainties in a digital soil mapping framework. **123**

4.1. Introduction **129**

4.2. Soil map to be validated **130**

4.2.1. Study area **130**

4.2.2. Digital soil mapping of the predictions **132**

4.2.3. Digital soil mapping of the uncertainties **134**

4.3. Materials and methods **136**

4.3.1. Concepts of soil map quality **136**
4.3.2. Probability sampling design **138**
4.3.3. Design-based estimation of soil map quality **141**
4.3.4. Conventional measures of map quality **142**
4.3.5. Implementation of methods **143**
4.4. Results and discussion **143**
4.5. Conclusions **149**
4.6. References **150**
Appendix 4.1. Accounting for measurement error in map validation **152**

Chapter 5. Snakes and ladders: manipulations of scale for digital soil mapping. **157**

5.1. Introduction **163**
5.2. Some concepts **164**
5.2.1. The digital soil map model **164**
5.2.2. Scaling digital soil maps **166**
5.3. Specific discussion of scaling methods for digital soil mapping **169**
5.3.1. Fine-gridding and coarse-gridding **169**
5.3.2. Deconvolution and convolution **173**
5.3.2.1. Convolution problems **173**
5.3.2.2. Block kriging example of P1 to P2 processes **175**
5.3.2.3. Further convolution problems **182**
5.3.2.4. Deconvolution problems **183**
5.3.3. Conflation and disseveration **184**
5.3.4. Further solutions for scaling problems **185**
5.4. Validation of soil information products generated from re-scaling methods **186**
5.5. Concluding remarks **188**
5.6. References **189**

Chapter 6. A general method for downscaling digital soil maps **191**

6.1. Introduction **197**

6.2. Materials and methods **199**
6.2.1. Algorithm for downscaling **199**
6.2.2. Downscaling using *dissever* **201**
6.2.3. Case study **203**
6.3. Results **204**
6.4. Discussion **208**
6.5. Conclusions **210**
6.6. References **212**

Chapter 7. General discussion, conclusions and future research 215

7.1. General discussion **219**
7.1.1. Comprehensive spatial soil information **219**
7.1.2. Practicable methodologies **220**
7.2. Summaries of research findings **224**
7.2.1. Mapping continuous depth functions **224**
7.2.2. Analysing uncertainties and the validation of digital soil maps **225**
7.2.3. Comprehensive digital soil mapping using low-cost input data **227**
7.2.4. Scale manipulation **229**
7.3. Overall research conclusions **230**
7.4. Future work **232**
7.5. References **235**

List of figures

Chapter 2.

Figure 2.2.1. The Edgeroi study area. Inset image is a true colour satellite image with the sampling locations superimposed as white dots. **51**

Figure 2.3.1. Available water capacity neural network prediction vs. observed plots from 80 randomly selected validation points at a) 0-10cm b) 30-40cm c) 80-100cm. Semi-variogram models of the residuals at each prediction depth (d-f). Final prediction (model prediction + residual) vs. observed plots at each prediction depth (g-i). RMSE: root mean square error. CCC: Lin's concordance correlation co-efficient. **60**

Figure 2.3.2. Validation of measured soil attribute (McGarry *et al.* 1989) vs. final predictions at 0–10cm, 10–20cm, 30–40cm, 70–80cm for a) carbon and b) AWC. CCC: Lin's concordance correlation co-efficient. **61**

Figure 2.3.3. Fitted splines (dashed lines) of observed carbon profile data (polygons) at five randomly selected sites (a-e). Digital soil map prediction depth functions of carbon (dashed lines) and observed carbon profile data (polygons) at same selected sites (f-j). **63**

Figure 2.3.4. Fitted splines (dashed lines) of observed AWC profile data (polygons) at five randomly selected sites (a-e). Digital soil map prediction depth functions of AWC (dashed lines) and observed AWC profile data (polygons) at same selected sites (f-j). **64**

Figure 2.3.5. Predicted total carbon (kg m^{-2}) and available water capacity (mm) to a depth of 1m across the Edgeroi study area. **66**

Figure 2.3.6. Predicted soil profile carbon (kg/m^3) to 1m displayed in 8 profile layers. **67**

Figure 2.3.7. Predicted soil profile available water capacity (m/m) to 1m displayed in 8 profile layers. **68**

Figure 2.3.8. Maps of scenario-based queries. A) Depth at which soil carbon decreases to below 1%. B) Depth at which cumulative total of soil carbon equals 5kg m^{-2}. C) Depth at which cumulative sum of AWC equals 100m. **71**

Chapter 3.

Figure 3.2.1. The prediction interval, characteristic features and general descriptive terminology. Adapted from Shrestha and Solomatine (2006). **91**

Figure 3.2.2. Flow diagram of the general procedure for achieving the outcome of mapping predictions and their uncertainties (upper and lower prediction limits) within a digital soil mapping framework. The 3 components for achieving this outcome are the prediction model, the empirical uncertainty model and the mapping component. **96**

Figure 3.4.1. Model validation of OC. Observed vs. fitted plots at 0–10cm, 30–40cm and 80–100cm before adding residuals (a-c) and after adding residuals (d-f). **104**

Figure 3.4.2. Box plots of the empirical distributions of model error as derived from the empirical uncertainty model at the selected depths of 0–10cm, 40–50cm and 80–100cm for AWC (a-c) and OC (d-f). **107**

Figure 3.4.3. Profile plots of AWC (a-c) and OC (d-f) at randomly selected validation sites. Bars represent actual observed values. Dotted lines represent final DSM predictions. Solid lines represent upper and lower prediction limits. **109**

Figure 3.4.4. Prediction interval coverage probability plots (PICPs) for OC (a) and AWC (b). **110**

Figure 3.4.5. Spatial variation of the degree of membership each instance has to each cluster including the extragrade class. Extragrade (a), cluster A (b), cluster B (c), cluster C (d), cluster D (f), cluster E (g). **112**

Figure 3.4.6. Variability of OC at 0–10cm, 30–40cm and 80–100cm across the Edgeroi study area. Lower prediction limit (1), DSM final prediction (2), upper prediction limit (3).

114

Figure 3.4.7. Variability of AWC at 0–10cm, 30–40cm and 80–100cm across the Edgeroi study area. Lower prediction limit (1), DSM final prediction (2), upper prediction limit (3).

115

Figure 3.4.8. Total water to 1m (a) and total OC to 1m (b) across the Edgeroi study area. Lower prediction limit (1), DSM final prediction (2), upper prediction limit (3). **117**

Chapter 4.

Figure 4.2.1. Lower Hunter Valley study area with respect to location in New South Wales (large box) and Australia (small box). **132**

Figure 4.2.2. Soil pH map to be validated (displaying only the 0-10 cm interval). Soil map depicts the (b) digital soil mapping prediction and the (a) upper and (c) lower prediction limits which constitute a prediction interval. **134**

Figure 4.3.1. Determination of sampling strata and their subsequent spatial extent. (a) Plot illustrating the process for constructing equal-area strata where the stratification variables were the depth-averaged whole-profile pH prediction and depth-averaged whole-profile difference between upper and lower prediction limits. Black lines indicate the threshold values for demarcation of each stratum A, B, C and D. (b) Spatial coverage of the equal-area strata across the study area. **140**

Figure 4.4.1. Plots of the observed soil pH vs. the corresponding digital soil mapped prediction of soil pH and resultant Lin's Concordance Correlation Coefficient at (a) 0-5 cm, (b) 5-15 cm, (c) 15-30 cm, (d) 30-60 cm, and (e) 60-100 cm. **145**

Figure 4.4.2. Prediction interval coverage probability plots for the areal proportion correctly predicted at (a) 0-5 cm, (b) 5-15 cm, (c) 15-30 cm, (d) 30-60 cm, (e) 60-100 cm. **147**

Chapter 5.

Figure 5.2.1. Generic soil map model (adapted from Bishop et al. 2001). Support of predictions are point when block *B* is very small. Block support prediction occur when *B* is greater than 0. Block support 1 is when *B* equal grid spacing *G*. Block support 2 is when *B* is greater than *G*. *B* may be larger than grid spacing. **165**

Figure 5.2.2. Exemplar soil map models. Panel 1 and Panel 2 have the same grid spacing yet Panel 2 is on block support (where block size is equal to the grid spacing), Panel 1 is on point support. Similarly for Panel 3 and Panel 4 except the grid spacing is larger. **167**

Figure 5.3.1. Panel 1- 5m point support thorium concentration (ppm) map. Panel 2- Associated variances of the measurement errors (also at 5m point support) of thorium concentration (ppm). **176**

Figure 5.3.2. Block support maps of thorium (ppm) where support size equals grid cell size (resolution) - a) 20m, b) 50m, and c) 80m. Panel 1- 'true' blocks created directly from 5m point support map with a P1 to P4 process- weighted averaging. Panel 2- P1 to P2 process- Block kriging with uncertain data. Panel 3- P1 to P2 process- Block kriging without including uncertainties. **178**

Figure 5.3.3. Map comparisons. Comparisons between 'true' block maps with map from block kriging with uncertain data a) 20m, b) 50m, and c) 80m. Comparisons between 'true' block maps with map from block kriging d) 20m, e) 50m, and f) 80m.

181

Chapter 6.

Figure 6.2.1. The downscaling algorithm written into the *dissever* program. **201**

Figure 6.3.1. Top Panel: SOC map displaying the variation of SOC in the top 0-30cm across the Edgeroi study area produced from the regression kriging procedure using observed soil data and a suite of environmental covariates. Middle Panel: Upscaled map of the same target variable with 1km by 1km blocks centred onto a 1km grid produced by block averaging. Bottom Panel: Map of the standard errors of predictions resulting from the block averaging procedure. **205**

Figure 6.3.2. Top Panel: Downscaled SOC map created from *dissever*. Middle Panel: Map of the absolute differences (given as two classes of difference) between the downscaled map (90m blocked map) and the 90m base map. Bottom Panel: Concordance plot between the 90m blocked map and the 90m base map. **207**

List of tables

Chapter 2.

Table 2.2.1. Accuracy of constructed neural networks for AWC with respect to iterative changes to the number of nodes. **55**

Chapter 4.

Table 4.3.1. Threshold values determined empirically of the stratification variables:- depth-averaged whole-profile prediction of pH and depth-averaged whole-profile difference between the upper and lower prediction limits for each stratum. **139**

Table 4.4.1. Results of the proposed soil map quality indicators: The areal proportion of the map within the specified prediction interval or correctly predicted (APCP) and Root mean square error of simulation (RMSES) at each depth increment. Additionally, corresponding measures of accuracy (RMSE), bias (ME) and imprecision (IMP) of the given map taking into account only the quality of the predictions. **144**

Chapter 5.

Table 5.2.1. Coordinate table of scaling processes based on attributes of existing map and scale attributes of desired map. **169**


users
require
assessing
water
services
prediction
models
making
project
DSM
quantifications
outputs
Develop
optimally
scale
Practicable
soils
many
practicable
sparse
explicit
assumptions
expertise
scales
property
something
need
tailored
whole-profile
Soil
functions
computational
whole
Furthermore
framework
critical
data
area
difficult
may
also
systems
time
population
quality
distribution
required
uncertainties
given
particular
methodologies
locations
observed
techniques
Methodologies
addressing
support
mapping
digital
generate
nutrients
derive
example
properties
predictions
quantitative
modelling
set
carbon
often
density
general
resources
sub-optimal
raster
statistical
quantifying
provide
within
method
comprehensive
range
available
spatial
information
methods
includes
human
include
means

General introduction

Soils and the functions they perform have and will probably always underpin human existence. This is because humans derive many direct and indirect services from soils, some of which include: (1) the production of food and fibre; (2) harmonising and regulating the cycling of nutrients; for example, carbon, nitrogen and phosphorous (Carton and Jarvis 2001; Loveland and Webb 2003); (3) playing a critical role in gas and climate regulation (Costanza et al. 1997); (4) storing of water, regulation of water quality and supply; (5) storage of biodiversity and genetic resources; (6) provisioning of cultural services; (7) provisioning of foundations for housing, transportation, and engineering structures.

However, humans are a dominating force on Earth (Crutzen 2002), where currently between 39-50% of the land surface has been transformed or degraded as a result of human activities over the past 300 years (Vitousek et al. 1997). In many parts, for example Africa where agriculture is practised without adequate fertilisers, the soils have been stripped of vital nutrients to support plant growth (Sanchez 2010); the critical function of soils providing food and fibre are threatened as a result. Degradation of soils through erosion, salinity and pollution mean that soils can not optimally provide the functions to which they are best suited. With a global human population expected to reach 10 billion by the end of this century (Crutzen 2002), the ability of the soils to optimally perform their functions will continually be more threatened.

Objective decision making is required for addressing soil degradation; for modelling, monitoring and measurement of particular soil services; and for general management of soil resources. There is clearly a need for relevant spatial soil information to complement these efforts. Spatial soil information denotes maps and associated databases which provide explicit, quantitative expressions of soil property variation for a given area. Soil properties may include organic carbon content, soil pH, and soil texture among others as well as functional soil properties such as available water capacity and carbon density. This information may be used directly for assessing particular soil functions or used indirectly as a proxy to predict more difficult-to-measure soil functions. Whatever the case, the audience of soil information users is not limited to soil scientists but includes climatologists, ecologists, hydrologists, geologists, and engineers amongst others who require

thematic (soil property) spatial soil information at a range of spatial scales, and on different supports (areal or point). Typical scales or resolutions at which spatial soil information might be needed are < 20m, 20m -2km, or >2km to respectively address local/field, regional/catchment, and national/continental/global questions (McBratney et al. 2000). The key objective of this project therefore is to set forth practicable methodologies for delivering comprehensive spatial soil information that is tailored to the requirements of the information users; as opposed to the common situation of soil information users making do with that which is available.

Comprehensive spatial soil information

The creation and population of spatial soil information systems can be achieved by Digital Soil Mapping (DSM; McBratney et al. 2003). DSM is a quantitative framework for prediction of soil properties (and soil classes), which is rooted in Jenny's soil factorial equation (Jenny 1941); the basis of the soil landscape modelling paradigm (Hudson 1992). DSM entails exploiting the availability of digital information that describes the environment (in other words, soil formation factors) to generate predictive models based on observed data recorded at sparse locations to derive inference at unvisited locations. Digital soil maps are the visual product and are displayed as a gridded raster image. The raster also acts a storage format of a densely populated spatially explicit soil information database.

Comprehensive soil information systems require more than just predictions of soil properties made within a DSM framework. The Encarta English dictionary defines comprehensive as *something that includes many details or aspects*. Synonyms for comprehensive include: across-the-board, extensive, widespread, wide-ranging and all-embracing. With this terminology we might expect comprehensive spatial soil information to be tailored accordingly to the scale and support for optimally addressing a particular question and to include estimates of uncertainties associated with the predictions. We might also expect to have some metric to describe how reliable the predictions *and* the quantifications of the uncertainties are. Furthermore comprehensive spatial soil information systems should also describe soil variability in terms of the whole soil profile or in other words the 3-dimensional distribution of soil properties. Advances in DSM are required and while techniques of mapping the whole profile distribution of soil properties is a relatively new area of research (McBratney et al. 2011), methodologies for

quantifying the uncertainties in addition to techniques for assessing the quality of the whole-profile predictions *and* their uncertainties at a range of scales have not really been explicitly developed before. This project will set forth some practicable methodologies for achieving these objectives.

Practicable methodologies

Practicable is defined as *something that is doable and capable of being carried out or put into effect*. Practicable methodologies need to jointly consider (1) the availability and quality of soil information for producing spatial soil information systems; (2) the expertise of the practitioners producing the spatial soil information; (3) the validity of modelling techniques and adherence to statistical assumptions; (4) and the computational effort required to run models.

The soil information required for DSM often involves using legacy soil information collected during conventional soil survey. This information is often sparse and more often than not does not constitute a proper statistical sample of the region to be mapped. The soil information itself ranges in quality and may contain uncertainties from various sources which together means this data is often sub-optimal for implementing advanced statistical methods. Furthermore, implementation of complex statistical models for example, Markov Chain Monte Carlo (MCMC) methods (Minasny et al. 2011), conditional stochastic simulations (Goovaerts 1997), or Residual Maximum Likelihood- Empirical Best Linear Unbiased Predictor (REML-EBLUP; Lark et al. 2006) may often be prohibitive because strict statistical assumptions need to be adhered and they also require significant computational resources when working with large datasets. This makes mapping at fine resolutions or at global, continental and national scales a difficult task and not at all practicable in the sense of the time, efforts and expertise required.

DSM has now moved on from a research phase to one where it is operational and is being used to address real world problems that require relevant soil information to address them (Sanchez et al. 2009; Grunwald et al. 2011). This does not mean that practicable methodologies are sub-optimal; rather they are general and can be implemented more-or-less easily given some training and a moderate level of expertise. These general methods can handle a wide range data types in terms of quality, are applicable for a number of soil properties and generate desirable outputs

regardless of the density of observed legacy soil data. Furthermore, the time to generate outputs should not be a burden to prohibit their use.

The aims of this project are to:

1. Develop a general method for digital soil mapping of the whole-profile distribution of soil properties
2. Develop an empirical method for quantifying the soil property prediction uncertainties within a digital soil mapping framework
3. Develop a simple method and additional criteria for validating whole-profile predictions of soil properties *and* the quantifications of their uncertainties.
4. Develop a systematic framework for scale manipulations or upscaling and downscaling techniques for digital soil mapping as a means of generating soil information products tailored to the needs of soil information end-users.

References

Carton, O.T. and Jarvis, S.C., 2001. Nitrogen and phosphorous cycles in agriculture. In: P. de Clercq et al. (Editors), Nutrient Management Legislation in European Countries. Wageningen Press, Wageningen, pp. 347.

Costanza, R., dArge, R., deGroot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., Oneill, R.V., Paruelo, J., Raskin, R.G., Sutton, P. and vanden Belt, M., 1997. The value of the world's ecosystem services and natural capital. Nature, 387: 253-260.

Crutzen, P.J., 2002. Geology of mankind. Nature, 415: 23-23.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Grunwald, S., Thompson, J.A. and Boettinger, J.L., 2011. Digital soil mapping and modelling at continental scales: finding solutions for global issues. Soil Science Society of America Journal, 75: 1201-1213.

Hudson, B.D., 1992. The soil survey as paradigm-based science. Soil Science Society of America Journal, 56: 836-841.

Jenny, H., 1941. Factors of Soil Formation, a System of Quantitative Pedology. McGraw-Hill, New York.

Lark, R.M., Cullis, B.R. and Welham, S.J., 2006. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. European Journal of Soil Science, 57: 787-799.

Loveland, P. and Webb, J., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. Soil & Tillage Research, 70: 1-18.

McBratney, A.B., Minasny, B., MacMillan, R.A. and Carre, F., 2011. Digital soil mapping. In: P.M. Huang, Y. Li and M.E. Sumner (Editors), Handbook of Soil Sciences: Properties and Processes. CRC Press, Boca Raton, FL, pp. 37:1-43.

McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S. and Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma, 97: 293-327.

McBratney, A.B., Santos, M.L.M. and Minasny, B., 2003. On digital soil mapping. Geoderma, 117: 3-52.

Minasny, B., Vrugt, J.A. and McBratney, A.B., 2011. Confronting uncertainty in model-based geostatistics using Markov Chain Monte Carlo simulation. Geoderma, 163: 150-162.

Sanchez, P.A., 2010. Tripling crop yields in tropical Africa. Nature Geoscience, 3: 299-300.

Sanchez, P.A., Ahamed, S., Carre, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonca-Santos, M.D., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vagen, T.G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A. and Zhang, G.L., 2009. Digital soil map of the world. Science, 325: 680-681.

Vitousek, P.M., Mooney, H.A., Lubchenco, J. and Melillo, J.M., 1997. Human domination of Earth's ecosystems. Science, 277: 494-499.


soil
mapping
model
information
soils
maps
map
spatial
properties
scale
data
digital
models
function
method
approach
prediction
methods
used
may
predictions
kriging
DSM
uncertainty
depth
variation
environmental
distribution
survey
observations
attributes
framework
classes
scales
century
point
based
one
variability
developed
land
often
fuzzy
development
geostatistics
describing
secondary
represents
variables
uncertainties
different
across
time
general
discussed
horizons
parameters
approaches
area
random
well
conventional
early
described
application
also
describe
support
attribute
region
validation
property
required
quantitative
continuous
classification
analysis
Soil
using
example
regression
domain
theory
factors
defined
knowledge
requires
predicted
observed
need
within
functions
modelling
carbon
sets
locations
set
quality
error
SSPFs

Soil mapping is one of the pillars to the challenge of sustainable development.

[Jeffrey Sachs 2009]

Chapter 1

Context of research: a review of the literature

1.1. Introduction

There is a well documented history of human interaction with soils that dates back to very early civilisation with the move from nomadic lifestyles to sedentary agriculture (Brevik and Hartemink 2010). Early humans saw soils as a provider of food and fibre and in the beginning, likely used a trial-and-error approach to determine the best places to farm. Even at this early stage of human civilisation much thought and possibly wonderment would have been given to the seemingly erratic variation of soils from one place to another. The rudimentary knowledge gained from observing these variations in soils possibly began the documentation and description of soils with the intention of describing and delineating those that were beneficial or detrimental for agriculture.

In present times our fascination with soils has not waned. In fact we view soils as providing much more than just our food and fibre and hence place probably more value on them for sustaining life and ecosystem functioning (Costanza 1997). Documenting and describing how soils vary across the Earth's land surface has advanced significantly from those rudimentary beginnings with many developed countries now having detailed inventories of soil resources at relatively fine scales (Cook et al. 2008). Even on a global scale, there are coarse representations that describe how soils vary from one place to another for the entire globe (Grunwald et al. 2011). The documentation of how soil varies generally comes in the form of maps, soil survey reports and recorded laboratory and field descriptions.

This review will document the role that soil inventory and soil mapping has had for complementing the management of soil resources. As it stands now, soil mapping is essential for planning and management, and makes decision making an objective process. The timeline begins from early history to the development of soil survey programs that began in the 19th century to new developments in soil mapping particularly in the digital form. Soil maps have evolved from representing soil variability in finite circumscribed regions as discrete tessellations of soil classes- which constitute a general-purpose classification; through to continuous

representations of soil property variation with associated uncertainties at all manner of spatial scales. This progression has been facilitated by the introduction of computers and advanced modelling techniques. Furthermore, an awareness of soils as critical for performing life-sustaining and ecosystem functions has seen an increased need for more detailed and relevant soil information; more advanced than what conventional soil maps can provide. The review will also address areas of soil mapping that require advancement in order to meet the increasingly sophisticated demands for comprehensive, relevant and tailored soil information. The subsequent chapters of this project will then explicitly detail how we might go about doing this with a practicable suite of methodologies.

1.2. Early mapping of soils

Much of the documented early history of soil mapping is referenced from Brevik and Hartemink (2010) where there is more detailed discussion in the context of the historical development of soil science. The earliest known soil maps as we know them today originated during the early 18th century in Europe where individual landholdings would prepare maps with notations such as fields for wheat, hemp or grapes. Krupenikov (1992) in Brevik and Hartemink (2010) detailed that land survey maps made in Russia in the 1760s reported on the quality of soils in various locations. During this time in Germany, soil mapping was also extensively used for taxation purposes and valuation of land. In fact, prior to the 18th century in Europe there was a general recognition of a direct relationship between soils and governments, where soils were believed to determine the economic vitality, governance, and national character of a country (Krupenikov 1992). The documentation of soils information in some form would have at least been necessary for those countries and regions interested in advancing their economic status.

Nineteenth century geologists and geological surveyors probably were the first to really initiate what is referred to now as the soil-landscape paradigm that states that soils are characterised as a function of parent material, climate, organisms, relief and time (Brown 2006, Hudson 1992). Early soil maps in the USA were made by state geologic surveys, for example the soil map of Massachusetts published in 1841 (Aldrich 1979 in Brevik and Hartemink 2010) and the 1882 soil map of Wisconsin (Coffey 1911 in Brevik and Hartemink 2010). The geologists however delineated

maps according only to geologic formation, lithology, and type of surficial deposits. Soil information was usually provided in an accompanying report and sometimes as independent soil maps (Brown 2006). The geologists were thus only able to realise the geologic and physiographic factors of parent material, topography and age for understanding and deciphering soil variability. It was not until the late 19th century that Russian naturalist V. V. Dokuchaev added the geographic factors of climate and vegetation to the geologic and physiographic factors. Dokuchaev's efforts virtually defined the framework of all subsequent soil mapping efforts from then on (Brown 2006). Jenny (1941) formalised these concepts with the derivation of the famous *clorpt* equation, intended as a mechanistic model for soil development:

$$S = f(cl, o, r, p, t)$$

where S stands for soil, cl represents climate, o represents organisms including humans, r represents relief, p represents parent material, and t represents time. Authors have argued Jenny's state factor model is virtually mathematically unsolvable (Huggett 1975). However, since this Russian-originated system of soil study has been developed, it has been used by numerous soil surveyors all over the world as a conceptual framework for understanding the factors that may be important for understanding soil variability across a region. It formed the basis of the national soil survey program in the United States which began in 1899 to assist frontier land development and planning (Simonson 1989). In the early to mid 20th century other countries followed with the development of detailed soil mapping programs to which Simonson (1989) describes with details.

1.3. Mapping soil spatial variation

Soil is often described as mantling the land more-or-less continuously with the exception being where there is bare rock and ice (Webster and Oliver 2006). As McBratney et al. (1992) describes, if we could measure all the properties of soil at all points in any region and then observe all the properties in their individual character spaces, there would each be a cloud of points with uneven density but without really sharp breaks. Meaning that, due to the complexity of soils and the complexity of landscape development, rapid changes in soil properties over short distances is generally an exception rather than a rule.

Nevertheless, our understanding of soil variation across any particular region derives from a small number of observations made in the field. Predictions are made at unobserved locations based on the properties of the soils that were observed at the specific locations to make a map of soil distribution across a mapping domain. There are two principal approaches for making predictions of soil at unobserved locations. The first of which divides the population of soil into more-or-less discrete classes. In the spatial context the classes are subdivisions of finite circumscribed regions of 'like' soils conforming to some general-purpose classification (Heuvelink and Webster 2001). The second approach treats soils as a suite of continuous variables that seeks to describe the way they vary across land (Heuvelink and Webster 2001). This approach is necessarily quantitative, as it requires numerical methods for interpolation between the locations of actual soil observations.

1.3.1. Conventional soil mapping

Because of its roots in geological survey and biological taxonomy, virtually all of the national soil survey programs carried out across the world during the 20th century mapped the distribution of soils using the discrete boundary approach. A number of authors including Simonson (1989), Arnold (2006) and Hewitt et al. (2008) have written exhaustively about the techniques and procedures required to make a soil map with this approach. It first entails direct observation of ancillary data (which includes aerial photos, geology, vegetation and topographic maps) and soil profile characteristics. Secondly, the observations of the soil characteristics are incorporated into a conceptual model that is used to infer soil variation. The conceptual model is primarily implicit which relies on the tacit knowledge of the soil surveyor (Hudson 1992). The third step involves applying this conceptual model across the survey area to predict at unobserved locations. Generally less than 0.001% of the survey region is actually observed (Burrough et al. 1971). The conceptual model is then transformed into a cartographic model (choropleth map), by drawing map unit boundaries on aerial photographs, showing the survey region tessellated into spatial classes (Scull et al. 2003). These classes constitute a general purpose classification that may be the classes of an established soil classification system or one based on experience and local landscape features (Webster 1977, Butler 1980).

1.3.2. Drawbacks of conventional soil mapping

Conventional soil mapping is based on the presumption that soil properties are homogeneous within each polygon and sharp breaks are assumed at the polygon boundaries (Heuvelink and Huisman 2000). Burrough et al. (1997) labelled this mapping concept the double-crisp model because the identified soil groups are supposed to be crisply delineated in both taxonomic space (the space defined by the soil properties) and geographic space (defined by the map unit boundary). Unfortunately this soil mapping concept is at odds with the continuous nature of soils and has often been questioned by numerous authors; examples include Webster and De La Cuanalo (1975), Nortcliff (1978) and Nettleton et al. (1991). As a model for describing soil variability, the double-crisp is inadequate because it ignores the continuous spatial variation of both soil-forming processes and the soils themselves. Another criticism levelled at conventional soil mapping is that it is conducted in a heuristic manner, requiring a dependence upon tacit knowledge of the soil surveyor that is generally unfalsifiable and therefore unverifiable in any objective sense (Hewitt 1993, Lagacherie et al. 1995).

For communicating and delivering comprehensive soil information to clients, sophisticated use of conventional soil maps is limited. Because of the implicit assumptions this approach has for explaining soil variability, the information is not suitable for quantitative studies; the language used to describe soils is abstract with a lot of technical jargon that are qualitative in nature (Hartemink et al 2008). The inaccuracies, imprecision and static nature of the maps are major drawbacks, as is the common situation of mapping scale which is seldom useful for addressing a particular question (Hartemink et al. 2008). The cost and time required for making soil maps is a slow and expensive process making it a prohibitive exercise to tailor maps of this kind to suit a targeted purpose. Furthermore, because the cartographic model uses polygons it is difficult to integrate this soil information with grid based forms of earth resource data (e.g. digital elevation models, satellite imagery and climate data).

1.3.3. Continuous and quantitative representations of soil variability

With the introduction of modern computers and development of statistical methods for soil science from around the 1960s there began an emphasis on treating the

descriptions of soil variability on a more continuous and quantitative basis (Heuvelink and Webster 2001).

1.3.3.1. Fuzzy sets

One of the advances was a generalisation of the double-crisp model which could be facilitated by using fuzzy sets. Fuzzy set theory (or fuzzy logic) was first developed by Zadeh (1965) as a method for allowing the matching of individuals to be determined on a continuous scale instead of a binary or nominal scale. Fuzzy sets are generalisations of discontinuous classes where the indicator function of crisp set theory, with values 0 or 1, is replaced by the membership function of fuzzy set theory, with values in the range 0 to 1 (McBratney and de Gruijter 1992). In other words, all individuals of a set share the properties defined for the set at certain degrees, which can range from 0 to 1. McBratney et al. (1992), McBratney and de Gruitjter (1992), Odeh et al. (1992), Lagacherie et al. (1997) and Grunwald et al. (2001) are a few notable examples which provide an overview of the applications of fuzzy sets in soil science. Examples of fuzzy sets for numerical classification and mapping are discussed as well as other applications such as for land evaluation, modelling and simulation of soil processes, fuzzy soil geostatistics, soil quality indices, and fuzzy measures of imprecisely defined soil phenomena. Though useful, fuzzy set theory have not been widely adopted for soil mapping (Grunwald and Lamsal 2006). One reason is possibly the of difficulty in interpreting the outputs. For example, instead of one map, there are potentially many for a single region. Furthermore, the unsupervised nature of the clustering algorithm and the requirement of sizeable datasets to generate meaningful results are additional reasons why fuzzy set theory has not been widely adopted in soil mapping.

1.3.3.2. Geospatial models

Soil classification systems and the practice of mapping soils have developed side by side since soil survey programs began in the early and mid 20th century across the world (Brown 2006). The language instilled in soil classifications and descriptions thereof have been suitable for communicating soils information amongst people with knowledge of soils, regardless of the flawed assumptions of conventional mapping. Nevertheless, communicating soils information to non-soil science minded people (but who have an interest and need for soils information) has proven difficult because

(1) of abstract technical jargon, and (2) they have a preference for soil information that describes the variability of soil properties across a survey region (Sanchez et al. 2009). Mapping the spatial variability of soil properties requires the use of geospatial prediction models as well as incorporating gridded or raster-based data models to make predictions onto and display maps (thus a trend away from the polygon data model).

Geospatial models predict soil attributes at un-sampled locations (such as regularly arranged grid cell nodes) largely by interpolation from observations distributed throughout a mapping region (McBratney et al. 2011). Soil attributes mean soil properties such as those including soil texture, soil organic carbon content, soil pH and other like properties. Soil attributes also includes soil classes, and as McBratney et al. (2003) reviewed, about 30% of studies that applied geospatial models for mapping soil attributes have predicted soil classes (70% predicted soil properties).

The first purely geospatial approaches were almost entirely based on geostatistics- a regionalised variable theory developed in the 1960s and 70s by French mathematician and geologist Georges Matheron. A precursor to geostatistics however was trend surfaces that required fitting some form of polynomial equation through soil (or environmental) attribute values (Grunwald 2006). Other purely geospatial models focused on modelling short-range (local) variations such as nearest neighbours, inverse distance weighting, and splines of which have been described in various forms by Laslett et al. (1987) and Burrough and McDonnell (1998). All these purely spatial approaches (i.e. they don't depend on any understanding of soil formation factors) for soil attribute mapping arose due to the fact that quantitative variables describing soil forming factors were not available (McBratney et al. 2011). This type of information was to come in the future however.

1.3.3.3. Geostatistics

Burgess and Webster (1980) were one of the first to use kriging, the practical application of geostatistics for soil attribute mapping. Many contributions detailing geostatistics and the underlying theory for soil scientists been written with notable works from Burrough (1993), Goovaerts (1999) and Webster and Oliver (2001).

Geostatistics provides a meaningful framework for mapping soil attributes because it treats observations as random variables. This concept works because soil

as it varies continuously, is the result of a complex interaction of physical, chemical, and biological processes that no description of it (soil) is complete or without error. Such variation of soils could be described as random, even chaotic (Webster 2000) and geostatistics provides the necessary tools that allow us to treat the variation of objects (in this case soil attributes) as random variables. By adopting the stochastic view, at each point within a mapping domain, there is not one value for an attribute, but a whole set of values. For example, at a location $x_{0,}$ a predicted soil attribute *S* is treated as a random variable with a mean (μ), variance (σ^2), and a cumulative distribution function (cdf) (Webster and Oliver 2001). Thus the set of random variables at each node in the mapping domain $S(x_1)$, $S(x_2), S(x_3) \ldots . S(x_n)$, constitutes one realisation of a random function or stochastic process.

Maintenance of the "law of geography" is integral to geostatistics which says that nearby things tend to be more alike than those far apart. This spatial auto-correlation is described by something called the variogram. Variograms measure the spatial auto-correlation of phenomena such as soil properties using semi-variance (McBratney and Pringle 1999). The average variance between any pair of sampling points (calculated as the semi-variance) for a soil property *S* at any vector of distance **h** apart can be given by the formula (Webster and Oliver 2001):

$$\gamma(h) = \frac{1}{2m\,(\mathbf{h})} \sum_{i=1}^{m\,(\mathbf{h})} \{s(\mathbf{x}_i) - s(\mathbf{x}_i + \mathbf{h})\}^2$$

[1.3.1]

where $\gamma(h)$ is the average semi-variance of the soil property, *m* is the number of pairs of sampling points *s* is the value of the property *S*, **x** is the coordinate of the point, and **h** is the lag (separation distance of point pairs). Thus in accordance with the "law of geography", in a 'normal' situation, points closer together show smaller semi-variance, whereas pairs of points farther away from each other should display larger semi-variance. A variogram is generated by plotting the average semi-variance against **h**. Various models can be fitted to the data to describe the variogram; four of the more common ones are the linear model, the spherical model, the exponential model, and the Gaussian model (Burrough 1993). Once an appropriate variogram has been modelled it is then used for distance weighted interpolation (kriging) at unvisited locations.

Since the initial use of geostatistics in the 1980s for soil mapping, numerous developments and refinements have extended the basic principles. One of the developments was co-kriging where early in the development of soil geostatistics it was recognised that soil could be better predicted if denser sets of secondary variables that were spatially cross-correlated with the primary soil attribute were available (see McBratney and Webster 1983 for an early example). Other forms of kriging have also evolved to either deal with non-normality (lognormal kriging, disjunctive kriging, indicator kriging), whereas others addressed varying trend or drift (universal kriging, kriging with external drift, regression kriging; Heuvelink and Webster 2001); all of which deal with continuous soil properties or classes, and can give estimates on support that have some defined area (blocks) or points. Moreover the application of geostatistical methods provides a quantitative measure of uncertainty through the kriging prediction error, which is useful for gauging the reliability and accuracy of the predicted map (Goovaerts 1999). Geostatistical developments have evolved with the development of Geographic Information System (GIS) technologies and with the availability of secondary data sets to describe the spatial distribution of soil forming factors. One major limitation of geostatistical approaches is the large amount of data required to define a meaningful model of spatial autocorrelation that is often not available. This is because the soil information used for these methods often comes from legacy observations carried out during soil survey for the original purposes of conventional soil mapping that collectively are often sparsely populated datasets (Burrough et al. 1971).

1.3.3.4. Multivariate geospatial models

The unification of concepts from Jenny's factorial soil-landscape model and purely geostatistical models was developed during the early 1990s. This represented an alternative spatial prediction strategy to map the distribution of soil attributes. Original applications of this work first began by using Jenny's factorial model in an explicit spatial prediction framework. While McBratney et al. (2011) cites earlier studies from the 1940s, 60s, and 70s as being precursors to the approach, it probably wasn't until the advent of GIS coupled with dissemination of gridded information sources which described soil forming factors (e.g. from digital elevation models, climate data and remote sensing imagery) did wide application of the approach really get under way.

The idea is that in order to make a map describing the distribution of a soil attribute across a given spatial domain, soil observation points are intersected with layers of secondary or environmental data; a model of some structure is fitted to describe the relationship between the soil observations and secondary data. This is followed by using the fitted model to predict at all locations (grid cell nodes of a raster) of the mapping domain. This approach is quantitative and relies on the necessary assumption that the soil observations are correlated with the secondary data. Subsequently, McKenzie and Austin (1993) termed the spatially explicit factorial approach as “environmental correlation”. Some of the first examples of this approach were conducted by Moore et al. (1993) who used terrain attributes derived from a DEM (15m resolution) to predict A-horizon thickness and pH for a region in Colorado, USA. Skidmore et al. (1991), Bell et al. (1992), Odeh et al. (1994), Lagacherie and Holmes (1997) and McKenzie and Ryan (1999) are well known pioneering studies of what will be referred to for now as the environmental correlation approach for predicting soil properties or classes.

It was during the mid 1990s that there came the realisation of a similarity between the geostatistical kriging approach and the environmental correlation approach and combined the two methods. The combinatorial approach became generically known as regression kriging (Odeh et al. 1995) which is a analogous to universal kriging (Stein and Corsten 1991), or kriging with external drift (Wackernagel 1998) or even kriging after de-trending (Goovaerts 1999). The difference between these geostatistical approaches with the generic regression kriging method is that they allow only for linear relationships to be derived between the soil attribute of interest and the secondary data. Nevertheless the approach entails fitting a model between the observed soil attribute data and the secondary data describing Jenny’s soil forming factors. Then kriging is performed on the residuals of the model fit, with the view that any spatial autocorrelation of the residuals can be captured as a means of improving the overall prediction of the target soil attribute. In practice the regression model prediction and the kriged residual are added together which terminates into a final prediction.

1.3.3.5. Digital soil mapping

These earlier geostatistical and environmental correlation approaches were generalised and formalised into the Digital Soil Mapping (DSM) framework

described by McBratney et al. (2003). Earlier work by Bishop et al. (2001) were probably the first to coin the term 'Digital Soil Mapping' however. Scull et al. (2003) also came up with a similar type of framework which they labelled 'Predictive Soil Mapping'. The DSM approach uses the *clorpt* formulation of Jenny (1941) not for explanation but for empirical quantitative descriptions of relationships between soils and spatially referenced environmental data, with a view of using these as soil spatial prediction functions (McBratney et al. 2011). This is called the "*scorpan*" model or the *scorpan*-SSPFe method, meaning a framework for soil mapping based on *scorpan* and Soil Spatial Prediction Functions (SSPFs) and spatially auto-correlated errors (*e*) (McBratney et al. 2003). The *scorpan* model is written as:

$$S_c[x, y, \sim t] \text{ or } S_p[x, y, \sim t]$$
$$= f(s[x, y, \sim t]\ , c[x, y, \sim t], o[x, y, \sim t], r[x, y, \sim t],\ p[x, y, \sim t]\ , a[x, y, \sim t], n) \quad \textbf{[1.3.2]}$$

where:

S_c= soil class

S_p= soil property

s = soils, other attributes of the soil at a point

c = climate, climatic properties of the environment at a point

o = organisms, vegetation, or fauna, or human activity

r = topography, landscape attributes

p = parent material, lithology

a = age, the time factor

n = space, spatial position

t = time (where *t* is defined as an approximate time)

x,y = the explicit spatial coordinates

f = function or soil spatial prediction function (SSPF)

Soil (*s*) is included as a factor because soil can be predicted from its properties, soil properties from its class or other properties (McBratney et al. 2003), and earlier studies have shown this, for example McBratney and Webster (1983). This additional soil information could be gathered from a prior soil map, or from either remote or proximal soil sensing or even expert knowledge. The *n* factor means that soil can be predicted from its spatial coordinates alone, but may also be used in a distance-to-or-

from some object, such as distance from a watercourse, distance from a road or proximity to a point source of pollution etc.

In the last 5-10 years, there has been a proliferation of high resolution environmental data that describe or can be used as proxies within the *scorpan* framework such as DEMs, and remote (satellite) and proximal sensing data infrastructures, much of which are discussed in McBratney et al. (2003) and Minasny et al. (2008). Subsequently there has been a significant uptake in using the *scorpan*-SSPF approach around the world as a means of building or populating spatial soil information systems from relatively sparse datasets (Lagacherie 2008). Updating or renewing existing or even out-of-date soil maps can be facilitated with DSM, see for example Rossiter (2006) and Kempen et al. (2009). In a more overall context, DSM is seen as a practicable framework for fulfilling the current and future demand for relevant soil information (Sanchez 2009). Authors have noted that DSM is now moving from the research phase to one that is operational in the context of global and national soil mapping ventures (Sanchez et al. 2009, Grunwald et al. 2011).

In practice, the creation of soil maps using the *scorpan*-SSPFe approach was discussed previously in the description of the environmental correlation and regression kriging approaches. Within some mapping domain, either existing or new observations *m* are taken from explicit locations [x, y]. This is followed by fitting some kind of function to a set of pedologically meaningful environmental layers which are generally gridded raster layers of a given spatial resolution. Once the model is fitted at the *m* observation points, the model is extended to all grid cell nodes of the raster layers, giving a digital soil map.

The form of the model or *f* (the SSPF) is often defined at the outset of a mapping project that is decided upon for reasons which may include: (1) it is one that the map producer is most familiar with; (2) its simplicity of application; (3) or its sophistication and power to distinguish complex relationships; (4) or that it captures some pedological nuance that the practitioner desires. Whatever the case, there are many forms of *f* that are available and the development of them continues with advances in normal statistical theory which are discussed generally in McBratney et al. (2000) and (2003) and reviewed extensively and in more detail by Hastie et al. (2009). SSPFs can range from simple linear models with either ordinary or generalised least squares fitting; generalised linear and additive models (Hastie and Tibshirani 1990); classification and regression tree models (Breiman et al. 1984);

neural networks (Hastie et al. 2009); advanced data mining methods such as Cubist models (Quinlan 1993; http://www.rulequest.com/); and knowledge-based inference systems (Zhu et al. 1997, Bui 2004). In terms of handling any spatial autocorrelation in the residual or *e* that is likely to result from fitting a *scorpan*-SSPF, a regression kriging (alternatively *scorpan* kriging; McBratney et al. 2011) or universal kriging methodology would be used.

Further developments to the DSM approach as a means of improving the accuracy of predictions has been the application of methods of generating multiple models via iteration which are then aggregated to produce final estimates. These approaches are discussed in McBratney et al (2003) and are referred to as bootstrap aggregating or bagging (Effron and Tibshirani 1993) and boosting methods (Freund and Schapire 1997). Bootstrapping is an iterative sampling (with replacement), model fitting method, for which is the basis of the powerful random forest algorithm (Brieman 2001) that has recently been popularised in DSM (Grimm et al. 2008 and Eustace et al. 2011 as notable examples). While boosting is a "committee" approach that combines the outputs of many "weak" models.

The other important development (in terms of strengthening model predictions) is the application of formal statistical methods for modelling spatial variation. REML-EBLUP (Residual Maximum Likelihood- Empirical Best Linear Unbiased Predictor) was introduced by Lark et al. (2006) as a statistically sound method for soil mapping. Intrinsically similar to regression kriging in that both are mixed models where the observed data are modelled as the additive combination of fixed effects (the secondary environmental data), random effects (the spatially correlated residuals *e*), and independent random error. Yet REML estimates the parameters of the trend and covariance functions unbiasedly. These parameters are then used in the EBLUP i.e. a general linear mixed model. The statistical theory of REML-EBLUP is discussed in Lark et al. (2006) and arose out of a need to rectify issues associated with normal regression kriging (the method proposed by Odeh et al. 1995) where estimation of the variogram of residuals is theoretically biased (Cressie 1993). While regression kriging may be theoretically sub-optimal and an *ad hoc* method, the improvement of prediction accuracy from REML-EBLUP has been demonstrated to be only small when comparative analysis was performed (Minasny and McBratney 2007). Because the REML estimation of parameters requires numerical optimization, the practicability of REML-EBLUP with large data sets has been questioned (Lark et al.

2006, Minasny and McBratney 2007). Nevertheless REML-EBLUP stands as a statistically sound soil mapping SSPF and Lark et al. (2006) provide discussion and ideas on improving the practical limitations this method may have.

1.4. Whole-profile digital soil mapping

Soil survey, soil classification and conventional mapping of soils consider soil as a 3-dimensional (3D) entity or body (Hole 1953). In the words (or to that effect) of American pioneering soil scientist Charles E. Kellogg: *"soils have shape and area, breadth and width, as well as depth"* (Kellogg 1949). However, DSM primarily has been used only in the 2D sense to predict soil property variation for single depth intervals or horizons- predominantly only from the top soil (Grunwald 2009). In fact, Grunwald (2009), who reviewed 90 journal articles from high impact soil science journals, found that 28% of reviewed studies performed *scorpan-SSPF* modelling for multiple soil layers or horizons; only one paper (Carrara et al. 2007) provided a 3D representation of soil which was of soil penetration resistance. DSM of soil classes is an exception for 3D soil mapping and has much utility because a number of soil properties can be inferred from one classification which are usually derived from the modal profile of each class. However the problem with mapping soil classes are that the soil properties vary discretely in the taxonomical space which may be problematic as soil variability across a mapping domain will appear as a stepped function rather than a fully continuous function (Webster and Oliver 2006).

It is quite a sensible undertaking to attempt to map the variation of soil properties in both the lateral and vertical dimensions. Understanding the carbon sequestration potential of soil and for carbon accounting (Bajtes 1996, Lal 2004), determining the amount of water soils can hold across a field or even a watershed, determining the depth to an impeding layer for crop growth across a farm, and investigations of soil acidity, (just to detail a few examples) will most likely require some understanding of how soil properties vary with depth. Ponce-Hernandez et al. (1986) describes that soil properties vary more-or-less continuously with depth. The variation is often anisotropic for certain properties such as carbon (Hiederer 2009) and soil texture (Myers et al. 2011), which may be the result of landuse activity or the gravitational vector of profile weathering and development or both (Hole 1961). Exceptions to continuous soil property variation with depth is where there is strong anthropogenic (cultivation, removal and replacement of soils), geologic (contrasting parent

materials), and pedological (the development of clear and abrupt soil horizons) forcings for which sharp discontinuities in the depth distribution of soil properties will occur.

Besides that of Carrara et al. (2007) a few studies have proceeded to map out the 3D distribution of soil properties, which have largely been GIS-based rather than model-based such as Mendonca Santos et al. (2000), Sirakov and Muge (2001) and Zhang et al. (2008). To model the distribution of soil properties in a pseudo-3D way requires more than just lumping model outputs for successive depth intervals or horizons together; rather some empirical function needs to be fitted to the observed soil data.

Empirical functions describing the depth distribution of soil properties include linear and polynomial functions (Colwell 1970; Moore et al. 1972), exponential and logarithmic functions (Russell and Moore 1968). Myers et al. (2011) introduced an asymmetric peak function for modelling complex and anisotropic soil property depth profiles with horizons of weathered loess. Smoothing splines for soil property variation were introduced by Erh (1972) which was followed by work from Ponce-Hernandez (1986) and mathematical derivation by Bishop et al. (1999) of pycnophylactic (mass preserving) smoothing splines. The mass preserving splines model the continuous variation of soil properties with depth whilst maintaining the integrity of the observed horizon or layer data. Bishop et al. (1999) demonstrated successfully the application of the mass-preserving spline for a number of soil properties with much success. The proviso for a 'good fit' is that a sufficient number of observations at regular depths are required.

A study by Minasny et al. (2006) used the negative exponential depth function to describe soil carbon concentration variation with depth in the Edgeroi area, Australia. The authors modelled the parameters of the exponential function using a modified neural network approach, then predicted parameters of the exponential function over the whole area, which enabled them to calculate the carbon distribution over the profile and also the storage of carbon at any depth. Mishra et al. (2009) fitted an exponential function to soil profile data from Indiana, USA, and then interpolated the parameters independently using ordinary kriging. Meersmans et al. (2009) performed something similar and developed empirical functions which predicted the parameters of the exponential depth function for the area of Flanders in Belgium. The functions were stratified based on landuse, and the parameters were related to particle size

distribution and height of groundwater. Kempen et al (2011) developed a depth function that combines general pedological knowledge with geostatistical modelling. They modelled the distribution of soil organic matter content based on typical horizons from ten soil types. Five depth function building blocks were defined, and for each soil type, the depth function structure was obtained by stacking a subset of modelled horizons. The parameters of the depth function for each of the horizons were interpolated using a geostatistical procedure combining environmental information.

One of the limitations for pseudo-3D DSM, particularly of using equivalents of the negative exponential depth function is that the function is only useful for certain soil properties such as soil organic carbon that naturally have that type of anisotropic variation. The method of Kempen et al. (2011) similarly suffers because it is suitable in circumstances such as where it was developed in the Netherlands where soil properties do not vary smoothly with depth because of anthropogenic and geologic forcings. A general method i.e. one that can be extended to a variety of soil properties and can cope to some extent with discontinuous variation in soil properties could be the use of the mass-preserving spline method of Bishop et al. (1999). To date this method has not been extended to a DSM framework, and as such is worthy of investigation (chapter 2).

1.5. Uncertainty estimation of digital soil maps

Implicit in DSM is the application of numerical models to infer the spatial (and temporal) variations of soil attributes from soil observation and also from knowledge of correlated environmental variables (Lagacherie and McBratney 2007). Unfortunately these models are not error free. This is because soil is so complex that no description or quantitative expression of them is ever complete (Webster and Oliver 2006); thus numerical models are merely abstractions of the true character and processes of soils. Quantitative methods for analysing the uncertainty of our representations of soil distribution are therefore indispensable in terms of communicating the reliability of given spatial soil information. Therefore it is not sufficient to produce digital soil maps showing the prediction of soil attributes only; the uncertainty of each prediction should also be shown (Minasny and Bishop 2008). However, possibly because practitioners are unaware of what to do in a relatively complex field, uncertainty is seldom quantified in routine DSM (Grunwald 2009).

Nevertheless, analysis of uncertainty is a major topic in many branches of science (May 2001), with some valuable contributions from Heuvelink (1998), Zhang and Goodchild (2002), and Foody and Atkinson (2002) whom provide substantive reviews of the theoretical aspects and practical applications of uncertainty analysis for use in the domain of spatial information sciences.

Minasny and McBratney (2002) describe the sources of uncertainty that might be present in soil information or may propagate through to the predictions when producing a digital soil map. These include: (1) uncertainties in the input data (the secondary environmental information) which are discussed in detail by Heuvelink (1998) and Bishop and Minasny (2006). (2) Uncertainties in the data used for calibrating SSPFs. These may be positional inaccuracies or measurement errors. (3) Uncertainties of the model parameters. (4) Uncertainties due to model structure.

A number of methods have been proposed to estimate uncertainty in the literature which Solomatine and Shrestha (2009) describe as falling into specific categories. These are: (1) analytical methods, one of which is first-order Taylor analysis (Heuvelink 1998). (2) Simulation and sampling-based methods, which includes the widely used Monte Carlo simulation method (Heuvelink 1998). (3) Bayesian methods (Diggle et al. 1998) which include Markov Chain Monte Carlo simulations as performed by Minasny et al. (2011), and Generalised Likelihood Uncertainty Estimation (GLUE; Beven and Binley 1992). (4) Methods based on analysis of model errors such as the empirical model error approach by Solomatine and Shrestha (2009). (5) Methods based on fuzzy set theory (Maskey et al. 2004).

First-order Taylor analysis is based on estimating the partial contribution of the error in each variable and evaluating its contribution to the overall uncertainty (Minasny and Bishop 2008). Categories 2 and 3 analyse the uncertainty of input variables (parameters, secondary environmental data,) by propagating it through the model (a SSPF) to the outputs, which requires some assumptions about the distributions of model parameters etc. Often, analysis of model errors (category 4) requires certain assumptions regarding the residuals (Solomatine and Shrestha 2009). The fuzzy theory based methods requires knowledge of the membership function of objects (soil attributes) subject to the uncertainty.

Probably one of the drawbacks to practicable application of uncertainty analysis is that most methods deal with only one source at a time (Solomatine and Shrestha 2009). Thus to determine the total uncertainty requires independent analyses. Not

only does this become a computationally laborious exercise, the contribution of uncertainty due to the interaction between each different source is not equated (Gupta et al. 2005). While it is important to know the contribution of each of the different sources of uncertainty, as then it becomes possible to address each in an objective manner, it is equally important to know the total prediction uncertainties. Practicable methodologies that are computationally efficient for quantifying the total uncertainty propagated through SSPFs are scarce and need to be investigated. Furthermore these practicable methodologies need to be extended in order to quantify the uncertainties of digital soil maps describing the distributions of soil properties both in the lateral and vertical dimensions (pseudo-3D). This is investigated in chapter 3.

1.6. Assessing the quality of digital soil maps with validation

As already established, SSPFs for mapping soils are rarely error free, and predictions are made in the presence of a number of sources of uncertainty. Validation of digital soil maps (the predictions) is analogous to ground-truthing exercises for remote sensing data infrastructures (Liu et al. 2007). It is a necessary, yet a rarely performed procedure (Grunwald 2009) for which some concept of soil map quality can be communicated. Quality criteria for digital soil maps in general are based on measures of variance between the SSPF predictions and co-located observations. For soil class maps the accuracy measure is often based on user's and producer's accuracies and the associated Kappa statistics which are discussed in detail by Lark (1995). For soil property maps, the Root Mean Square prediction Error (RMSE; for derivation see equation 2.2.3) is commonly reported.

Strategies for validating digital soil maps include: (1) holding back a random proportion of soil observations which are excluded from model calibration. Commonly around 25-30% of data are held back, which are then used to quantify the accuracy of the model. (2) Cross validation, which can be leave-one-out or *n*-fold cross-validation, and is similar to the random holdback strategy (Effron and Tibshirani, 1993). In leave-one-out cross validation (LOCV), a sample observation is left out, while the rest of the observations are used to calibrate the prediction model, the left-out observation is used to assess the accuracy of the calibrated model. The process is repeated for all observations. Meanwhile in *n*-fold cross validation, the dataset is divided into n sections or folds, and the cross validation process is repeated

for *n*-folds. (3) Additional samples are collected from the mapping domain and are completely independent of the model calibration data. Ideally, the additional samples are collected using a randomised or probability sampling design (de Gruijter et al. 2006). For random sampling, all units within a mapping domain have a positive probability of being selected, where the probabilities are determined by the sampling design and can be derived from this design (de Gruijter et al. 2006). In a review by Brus et al. (2011) they concluded that probability sampling is the more superior validation method (in comparison the random holdback or cross-validation strategies) because unbiased estimates of the measures of soil map quality can be obtained by 'design-based' inference and thus are free of model assumptions (Brus and de Gruijter, 1997). This is generally not the case for random holdback or cross-validation because the samples used for validation (and for calibration for that matter) do not represent a probability sample because the data often comes from legacy soil survey where sampling is performed without any statistical strategy (Brus et al., 2011). Thus the quality measures may be affected by bias.

The constraints of cost and size of soil datasets will determine the most appropriate validation strategy, such that random holdback may only be applicable for large datasets for example. Nevertheless, validation of soil maps has conventionally been geared only towards assessing the accuracy of the model predictions. Ideally, however, the quality combination of the predictions *and* the quantifications of the associated uncertainties need to be tested to properly evaluate digital soil maps (McBratney et al. 2011). An example of situations where this has been done is scarce, yet is the focus of investigations in chapter 4.

1.7. Serving the end user: scale and manipulations thereof for digital soil mapping

The discussion of DSM till now has detailed only general concepts of creating digital soil maps and made considerations for estimating soil distribution as a 3D entity within a DSM framework. Discussion then considered various aspects of analysing the prediction uncertainties and validating soil maps. Equally important to this discussion also is the considerations of scale. How do we represent scale with digital soil maps? At what scales are users likely to want spatial soil information? How does scale contribute to our descriptions and interpretations of soils? How does scale determine how and what methods are used to map soils? How do we go about

manipulating the scale of an existing soil map to create a new soil map at a desired scale?

When considering the hierarchy of spatial scales recognised for soil (Hoosbeek and Bryant 1992), the *i-levels* of interest for soil mapping would be global (*i*+6), continental (*i*+5), regional (*i*+4), watershed (*i*+3), farm (*i*+2) and field (*i*+1) scales. Conventional soil mapping uses scale in the cartographic sense, where the scale of the map indicates how the size of the paper map relates to the actual size of the area being portrayed. Generally low levels of the scale hierarchy are characterised by smaller spatial scales (areal extents) with lots of detail. At a field scale, a 1:500 scale would not be uncommon, where 1cm on the map would represent 5m on the ground. At the global scale, we would expect large spatial scales and less detail. For example, the FAO-UNESCO Soil Map of the World (Nachtergaele 1999) was produced at 1: 5 000 000 scale (1 cm on map is 50km on the ground).

Digital soil maps are produced using a gridded raster format where representations of scale are mostly defined in terms of the raster model resolution (Hengl et al. 2006). The other important entity of scale to consider for DSM is the matter of support, such as whether the predictions represent the value for a definable area or volume (referred to spatially as a block) or whether they represent the value at a point (very small area or volume) (Bierkens et al. 2000). Like conventional soil maps, the spatial extent of mapping is used to characterise scale in some sense for the different hierarchical levels. Extent, resolution, and support thus represent a scaling triplet, and constitute the digital soil map model which is described more-or-less by Bishop et al. (2001).

McBratney et al. (2000) indicated that the typical extent for global and continental scales would be >200 km with a typical resolution of >2 km. The *GlobalSoilMap.net* project is an exception to this, which is aiming to map a number of functional soil properties at a global scale with a resolution of 90m (Sanchez et al 2009). However, the intention of the *GlobalSoilMap.net* project is to provide soil information for the entire globe that can be used to address questions at all the meaningful scales in the hierarchy for soil mapping. Nevertheless, at the global scale we might expect that DSM might be used to address questions related to global climate modelling and policy making for ensuring food security (McBratney et al. 2000). At the regional and watershed scales, which will typically have extents of 2-200 km and resolutions of 20 m-2 km, the questions might be framed on monitoring

the environment in terms of soil degradation or studying changes brought about by disruption of ecosystems or specific landscapes. At the farm and field scales questions related to nutrient cycling might be addressed or alternatively the fine scaled soil information (typical extent of 2 km and resolutions of <20 m) could be used as input for decision making within a precision agriculture setting.

Important considerations for DSM at the meaningful scales of the hierarchy are the scale dependent soil-landscape relationships, which is often a neglected area of investigation (McBratney et al. 2011). The *clorpt* soil-forming factors all operate at different scales and therefore influence soil processes and soil variation at different scales as well (Addiscott 1993). Subsequently, as Heuvelink and Pebesma (1999) described, it is normal to expect different models (SSPFs) describing the same process (such as soil property variation) at different scales. A reason for this could be the relative importance of the environmental data layers at the investigated scale. Bui et al. (2006), Smith et al. (2006) and Kim and Zheng (2011) are examples which have investigated the scale influence of the environmental data (emphasis on digital elevation models) on the predictability of SSPFs. From a pedological perspective, a need to identify which scales of variation, which are operative and discernable in the environmental layers, are most strongly related to the observed variation in the soil attribute of interest represents good DSM practice (McBratney et al. 2011). However, there is a limitation to this such that often the scale at which environmental data are available is not at the scale required for investigation, which may lead to a simplification of SSPFs (Heuvelink and Pebesma 1999). Another reason for ending up with different SSPFs at different scales relates to the support, whereby SSPFs developed at point support do not necessarily extend to block support. Moving from point support to block support usually implies going from single point predictions to predictions which represent an average of point values within each block (Heuvelink and Pebesma 1999).

The last point of discussion is the matter of scale manipulation. Scale manipulation involves the practice of harmonising the scale at which soil information is available to the scale at which it is required. The practice of harmonisation represents a critical tool for delivering tailored soil information to the parties whom request it. For example, a map displaying the distribution of soil organic carbon may have been created at the regional scale to address particular questions related to that scale. However, a soil information user may require a map of the distribution of

carbon across a particular farm within the regional mapping domain. The practitioner could use the regional map to perform scale manipulations to produce a map with the ideal extent, resolution and support required for the farm scale. This would entail some sort of downscaling. Conversely some sort of upscaling or convolution would be necessary if maps at farm scale are used as inputs to deliver a map at the watershed or regional scale.

In geostatistics scale manipulation as described here is referred to as the 'change-of-support' problem for combining incompatible datasets (Cressie 1996). Change of support represents a wide spectrum of statistical challenges and associated methodologies that have been reviewed extensively by Gotway and Young (2002). For DSM, manipulations of scale involves methods of moving up and down the scale hierarchy, which are generally referred to as upscaling and downscaling methods respectively (Bierkens et al. 2000). Changes of support such as from point support to block support and vice versa also come under the umbrella of scale manipulation. Some methods for scale manipulation of soil information are discussed in McBratney (1998) and Bierkens et al. (2000). A general assumption of most methods assumes that the behaviour of soil at large scales is explained by the average of the soil behaviour at the fine scales. This implies some sort of linear relationship both between the hierarchies of scale and the support of observations. This assertion may or may not be upheld in reality or may only be relevant at a specific range of scales. Grunwald et al (2011), citing deYoung et al. (2008), does explain however, that nonlinear dynamics and alternate states are well known in ecological systems, yet they have been poorly investigated in the soil science domain.

An explicit framework of how to up- and downscale within the constraints of the digital soil map model has not previously been introduced, yet constitutes the focus of investigations in chapter 5. A novel method of downscaling is also investigated in chapter 6.

1.8. Thesis outline

Tailoring spatial soil information to the requirements of the user is one of the idealistic pursuits of this project; with the aim that the methods for delivering this tailored information are practicable. This review revealed some areas of new research that need to be conducted if a framework is put into place for delivering

comprehensive spatial soils information to the wide audience of soil information users. These areas are:

1. There is a requirement for a general method for digital soil mapping for the whole-profile (effectively pseudo-3D) distribution of soil properties. (Chapter 2)
2. There is a requirement for a general method for quantifying the total prediction uncertainties of the digital soil maps that describe the whole-profile distribution of soil properties, and which is not computationally prohibitive. (Chapter 3)
3. There is a requirement for a method to validate the whole-profile predictions of soil properties *and* the quantifications of their uncertainties. (Chapter 4)
4. There is a requirement for a systematic framework for scale manipulate the scale of digital soil maps as a means of generating soil information products tailored to the needs of soil information users. (Chapter 5 and 6).

1.9. References

Addiscott, T.M., 1993. Simulation modelling and soil behaviour. Geoderma, 60: 15-40.

Arnold, R.W., 2006. Soil survey and soil classification. In: S. Grunwald (Editor), Environmental Soil Landscape Modelling: Geographic Information Technologies and Pedometrics. Taylor and Francis, Boca Raton, pp. 37-60.

Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. European Journal of Soil Science, 47: 151-163.

Bell, J.C., Cunningham, R.L. and Havens, M.W., 1992. Calibration and validation of a soil-landscape model for predicting soil drainage class. Soil Science Society of America Journal, 56: 1860-1866.

Beven, K. and Binley, A., 1992. The future of distributed models- model calibration and uncertainty prediction. Hydrological Processes, 6: 279-298.

Bierkens, M.F.P., Finke, P.A. and de Willigen, P., 2000. Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bishop, T.F.A., McBratney, A.B. and Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, 91: 27-45.

Bishop, T.F.A., McBratney, A.B. and Whelan, B.M., 2001. Measuring the quality of digital soil maps using information criteria. Geoderma, 103: 95-111.

Bishop, T.F.A. and Minasny, B., 2006. Digital soil-terrain modelling: the predictive potential and uncertainty. In: S. Grunwald (Editor), Environmental Soil Landscape Modelling: Geographic Information Technologies and Pedometrics. Taylor and Francis, Boca Raton.

Breiman, L., 2001. Random forests. Machine Learning, 45: 5-32.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., 1984. Classification and Regression Trees. Wadsworth, Belmont, CA.

Brevik, E.C. and Hartemink, A.E., 2010. Early soil knowledge and the birth and development of soil science. Catena, 83: 23-33.

Brown, D.J., 2006. A historical perspective on soil-landscape modelling. In: S. Grunwald (Editor), Environmental Soil Landscape Modelling: Geographic Information Technologies and Pedometrics. Taylor and Francis, Boca Raton, pp. 61-104.

Brus, D.J. and de Gruijter, J.J., 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma, 80: 1-44.

Brus, D.J., Kempen, B. and Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. European Journal of Soil Science, 62: 394-407.

Bui, E.N., 2004. Soil survey as a knowledge system. Geoderma, 120: 17-26.

Bui, E.N., Henderson, B.L. and Viergever, K., 2006. Knowledge discovery from models of soil properties developed through data mining. Ecological Modelling, 191: 431-446.

Burgess, T.M. and Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties. 1. The semi-variogram and punctual kriging. Journal of Soil Science, 31: 315-331.

Burrough, P.A., 1993. Soil variability: a late 20th century view. Soil and Fertilizers, 56: 529-562.

Burrough, P.A., Beckett, P.H.T. and Jarvis, M.G., 1971. Relation between cost and utility in soil survey. Journal of Soil Science, 22: 359-370.

Burrough, P.A. and McDonnell, R.A., 1998. Principles of Geographic Information Systems: Spatial Information Systems and Geostatistics. Oxford University Press, New York.

Burrough, P.A., van Gaans, P.F.M. and Hootsmans, R., 1997. Continuous classification in soil survey: Spatial correlation, confusion and boundaries. Geoderma, 77: 115-135.

Butler, B.E., 1980. Soil Classification for Soil Survey. Oxford University Press, Oxford.

Carrara, M., Castrignano, A., Comparetti, A., Febo, P. and Orlando, S., 2007. Mapping of penetrometer resistance in relation to tractor traffic using multivariate geostatistics. Geoderma, 142: 294-307.

Colwell, J.D., 1970. A statistical-chemical characterization of four great soil groups in southern New South Wales based on orthogonal polynomials. Australian Journal of Soil Research, 8: 221–238.

Cook, S.E., Jarvis, A. and Gonzalez, J.P., 2008. A new global demand for digital soil information. In: A.E. Hartemink, A.B. McBratney and M.L. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer Science, Dordrecht, pp. 31-42.

Costanza, R., dArge, R., deGroot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., Oneill, R.V., Paruelo, J., Raskin, R.G., Sutton, P. and vandenBelt, M., 1997. The value of the world's ecosystem services and natural capital. Nature, 387: 253-260.

Cressie, N., 1996. Change of support and the modifiable areal unit problem. Geographical Systems, 3: 159-180.

Cressie, N.A.C., 1993. Statistics for Spatial Data. John Wiley and Sons, New York.

de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P. and Knotters, M., 2006. Sampling for Natural Resource Monitoring. Springer-Verlag Berlin Heidelberg, The Netherlands.

Diggle, P.J., Tawn, J.A. and Moyeed, R.A., 1998. Model-based geostatistics. Journal of the Royal Statistical Society Series C-Applied Statistics, 47: 299-326.

Effron, B. and Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability 57. Chapman and Hall, London, UK.

Erh, K.T., 1972. Application of spline functions to soil science. Soil Science, 114: 333–338.

Eustace, A.H., Pringle, M.J., Denham, R.J., 2011. A risk map for gully locations in central Queensland, Australia. European Journal of Soil Science 62(3), 431-441.

Foody, G.M. and Atkinson, P.M., 2002. Uncertainty in Remote Sensing and GIS. J. Wiley, Hoboken, NJ.

Freund, Y. and Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55: 119-139.

Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma, 89: 1-45.

Gotway, C.A. and Young, L.J., 2002. Combining incompatible spatial data. Journal of the American Statistical Association, 97: 632-648.

Grimm, R., Behrens, T., Maerker, M. and Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island - Digital soil mapping using Random Forests analysis. Geoderma, 146: 102-113.

Grunwald, S., 2006. What do we really know about the space-time continuum of soil-landscapes? In: S. Grunwald (Editor), Environmental Soil Landscape Modelling: Geographic Information Technologies and Pedometrics. Taylor and Francis, Boca Raton, pp. 3-36.

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modelling approaches. Geoderma, 152: 195-207.

Grunwald, S. and Lamsal, S., 2006. The impact of emerging geographic information technology on soil-landscape modelling. In: S. Grunwald (Editor), Environmental Soil Landscape Modelling: Geographic Information Technologies and Pedometrics. Taylor and Francis, Boca Raton, pp. 127-154.

Grunwald, S., McSweeney, K., Rooney, D.J. and Lowery, B., 2001. Soil layer models created with profile cone penetrometer data. Geoderma, 103: 181-201.

Grunwald, S., Thompson, J.A. and Boettinger, J.L., 2011. Digital Soil Mapping and Modeling at Continental Scales: Finding Solutions for Global Issues. Soil Science Society of America Journal, 75: 1201-1213.

Gupta, H.V., Beven, K. and Wagener, T., 2005. Model calibration and uncertainty estimation. In: M.G. Anderson (Editor), Encyclopaedia of Hydrological Sciences. John Wiley, New York, pp. 2015-2031.

Hartemink, A.E., Hempel, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., MacMillan, R.A., Minasny, B., Montanarella, L., Mendonca-Santos, M.L., Sanchez, P., Walsh, M. and Zhang, G., 2008. GlobalSoilMap.net- a new digital soil map of the world. In: J.L. Boettinger, D.W. Howell, A.C. Moore, A.E. Hartemink and S. Kienast-Brown (Editors), Digital Soil Mapping: Bridging Research, Environmental Application, and Operation. Springer Science, Dordrecht.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York.

Hastie, T.J. and Tibshirani, R.J., 1990. Generalized Additive Models. Chapman and Hall, London, England.

Hengl, T., 2006. Finding the right pixel size. Computers & Geosciences, 32: 1283-1298.

Heuvelink, G.B.M., 1998. Error Propagation in Environmental Modelling. Taylor and Francis, London.

Heuvelink, G.B.M. and Huisman, J.A., 2000. Choosing between abrupt and gradual spatial variation? In: H.T. Mowrer and R.G. Congalton (Editors), Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing. Ann Arbor Press, Chelsea, MI, pp. 111-117.

Heuvelink, G.B.M. and Pebesma, E.J., 1999. Spatial aggregation and soil process modelling. Geoderma, 89: 47-65.

Heuvelink, G.B.M. and Webster, R., 2001. Modelling soil variation: past, present, and future. Geoderma, 100: 269-301.

Hewitt, A.E., 1993. Predictive modelling in soil survey. Soil and Fertilizers, 56: 305-314.

Hewitt, A.E., McKenzie, N.J., Grundy, M.J. and Slater, B.K., 2008. Qualitative survey. In: N.J. McKenzie, M.J. Grundy, R. Webster and A.J. Ringrose-Voase (Editors), Guidelines for Surveying Soil and Land Resources. CSIRO Publishing, Collingwood, Australia, pp. 285-306.

Hiederer, R., 2009. Distribution of organic carbon in soil profile data. EUR 23980 EN. Office for Official Publications of European Communities, Luxembourg.

Hole, F.D., 1953. Suggested terminology for describing soils as three-dimensional bodies. Proceedings. Soil Science Society of America, 17: 131-135.

Hole, F.D., 1961. A classification of pedoturbations and some other processes and factors of soil formation in relation to isotropism and anisotropism. Soil Science 91: 375-377.

Hoosbeek, M.R. and Bryant, R.B., 1992. Towards the quantitative modelling of pedogenesis- a review. Geoderma, 55: 183-210.

Hudson, B.D., 1992. The soil survey as paradigm-based science. Soil Science Society of America Journal, 56: 836-841.

Huggett, R.J., 1975. Soil landscape systems- model of soil genesis. Geoderma, 13: 1-22.

Jenny, H., 1941. Factors of Soil Formation, a System of Quantitative Pedology. McGraw-Hill, New York.

Kellogg, C.E., 1949. Soil classification- introduction. Soil Science, 67: 77-80.

Kempen, B., Brus, D.J., Heuvelink, G.B.M. and Stoorvogel, J.J., 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. Geoderma, 151: 311-326.

Kempen, B., Brus, D.J. and Stoorvogel, J.J., 2011. Three-dimensional mapping of soil organic matter content using soil type-specific depth functions. Geoderma, 162: 107-123.

Kim, D. and Zheng, Y.B., 2011. Scale-dependent predictability of DEM-based landform attributes for soil spatial variability in a coastal dune system. Geoderma, 164: 181-194.

Lagacherie, P., 2008. Digital soil mapping: a state of the art. In: A.E. Hartemink, A.B. McBratney and M.L. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 3–14.

Lagacherie, P., Cazemier, D.R., vanGaans, P.F.M. and Burrough, P.A., 1997. Fuzzy k-means clustering of fields in an elementary catchment and extrapolation to a larger area. Geoderma, 77: 197-216.

Lagacherie, P. and Holmes, S., 1997. Addressing geographical data errors in a classification tree for soil unit prediction. International Journal of Geographical Information Science, 11: 183-198.

Lagacherie, P., Legros, J.P. and Burrough, P.A., 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. Geoderma, 65: 283-301.

Lagacherie, P. and McBratney, A.B., 2007. Spatial soil information systems and spatial soil inference systems. In: P. Lagacherie, A.B. McBratney and M. Voltz (Editors), Digital Soil Mapping- An Introductory Perspective. Elsevier, Amsterdam, pp. 301-326.

Lal, R., 2004. Soil carbon sequestration to mitigate climate change. Geoderma, 123: 1-22.

Lark, R.M., 1995. Components of accuracy of maps with special reference to discriminant-analysis on remote sensor data. International Journal of Remote Sensing, 16: 1461-1480.

Lark, R.M., Cullis, B.R. and Welham, S.J., 2006. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. European Journal of Soil Science, 57: 787-799.

Laslett, G.M., McBratney, A.B., Pahl, P.J. and Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. Journal of Soil Science, 38: 325-341.

Liu, C., Frazier, P. and Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. Remote Sensing of Environment, 107: 606-616.

Maskey, S., Guinot, V. and Price, R.K., 2004. Treatment of precipitation uncertainty in rainfall-runoff modelling: a fuzzy set approach. Advances in Water Resources, 27: 889-898.

May, R., 2001. Risk and uncertainty. Nature, 411: 891-891.

McBratney, A.B., 1998. Some considerations on methods for spatially aggregating and disaggregating soil information. Nutrient Cycling in Agroecosystems, 50: 51-62.

McBratney, A.B. and de Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science, 43: 159-175.

McBratney, A.B., de Gruijter, J.J. and Brus, D.J., 1992. Spatial prediction and mapping of continuous soil classes. Geoderma, 54: 39-64.

McBratney, A.B., Mendonca-Santos, M.L. and Minasny, B., 2003. On digital soil mapping. Geoderma, 117: 3-52.

McBratney, A.B., Minasny, B., MacMillan, R.A. and Carre, F., 2011. Digital soil mapping. In: P.M. Huang, Y. Li and M.E. Sumner (Editors), Handbook of Soil Sciences: Properties and Processes. CRC Press, Boca Raton, FL, pp. 37:1-43.

McBratney, A.B., Pringle, M.J., 1999. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. Precision Agriculture 1, 125-152.

McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S. and Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma, 97: 293-327.

McBratney, A.B. and Webster, R., 1983. Optimal interpolation and isarithmic mapping of soil properties. 5. co-regionalisation and multiple sampling strategy. Journal of Soil Science, 34: 137-162.

McKenzie, N.J. and Austin, M.P., 1993. A quantitative Australian approach to medium and small-scale surveys based on soil stratigraphy and environmental correlation. Geoderma, 57: 329-355.

McKenzie, N.J. and Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma, 89: 67-94.

Meersmans, J., van Wesemael, B., De Ridder, F. and Van Molle, M., 2009. Modelling the three-dimensional spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma, 152: 43-52.

Mendonca-Santos, M.L., Guenat, C., Bouzelboudjen, M. and Golay, F., 2000. Three-dimensional GIS cartography applied to the study of the spatial variation of soil horizons in a Swiss floodplain. Geoderma, 97: 351-366.

Minasny, B. and Bishop, T.F.A., 2008. Analysing uncertainty. In: N.J. McKenzie, M.J. Grundy, R. Webster and A.J. Ringrose-Voase (Editors), Guidelines for Surveying Soil and Land Resources. CSIRO Publishing, Collingwood, Australia, pp. 383-391.

Minasny, B. and McBratney, A.B., 2002. Uncertainty analysis for pedotransfer functions. European Journal of Soil Science, 53: 417–429.

Minasny, B. and McBratney, A.B., 2007. Spatial prediction of soil properties using EBLUP with the Matern covariance function. Geoderma, 140: 324-336.

Minasny, B., McBratney, A.B. and Lark, R.M., 2008. Digital soil mapping technologies for countries with sparse data infrastructures. In: A.E. Hartemink, A.B. McBratney and M.D. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer, Australia, pp. 15-30.

Minasny, B., McBratney, A.B., Mendonca-Santos, M.L., Odeh, I.O.A. and Guyon, B., 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. Australian Journal of Soil Research, 44: 233-244.

Minasny, B., Vrugt, J.A. and McBratney, A.B., 2011. Confronting uncertainty in model-based geostatistics using Markov Chain Monte Carlo simulation. Geoderma, 163: 150-162.

Mishra, U., Lal, R., Slater, B., Calhoun, F., Liu, D. and Van Meirvenne, M., 2009. Predicting soil organic carbon stock using profile depth distribution functions and ordinary kriging. Soil Science Society of America Journal, 73: 614-621.

Moore, A.W., Russell, J.S. and Ward, W.T., 1972. Numerical analysis of soils: a comparison of three soil profile models with field classification. Journal of Soil Science, 23: 193–209.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Science Society of America Journal, 57: 443-452.

Myers, D.B., Kitchen, N.R., Sudduth, K.A., Miles, R.J., Sadler, E.J. and Grunwald, S., 2011. Peak functions for modelling high resolution soil profile data. Geoderma, 166: 74-83.

Nachtergaele, F.O., 1999. From the soil map of the world to the digital global soil and terrain database: 1960-2002. In: M.E. Sumner (Editor), Handbook of Soil Science. CRC Press, Boca Raton, pp. 5-17.

Nettleton, W.D., Brasher, B.R. and Borst, G., 1991. The taxadjunct problem. Soil Science Society of America Journal, 55: 421-427.

Nortcliff, S., 1978. Soil variability and reconnaissance soil mapping- statistical study in Norfolk. Journal of Soil Science, 29: 403-418.

Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1992. Soil pattern recognition with fuzzy c-means: application to classification and soil-landform inter-relationships. Soil Science Society of America Journal, 56: 505-516.

Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma, 63: 197-214.

Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes- heterotopic co-kriging and regression kriging. Geoderma, 67: 215-226.

Ponce-Hernandez, R., Marriott, F.H.C. and Beckett, P.H.T., 1986. An improved method for reconstructing a soil-profile from analysis of a small number of samples. Journal of Soil Science, 37: 455-467.

Quinlan, R., 1993. Combining instance-based and model-based learning. In: Utgoff (Editor), Machine Learning '93. Morgan Kaufmann, San Mateo, CA.

Rossiter, D.G., 2008. Digital soil mapping as a component of data renewal for areas with sparse soil data infrastructures. In: A.E. Hartemink, A.B. McBratney and M.D. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Soil Data. Springer, Australia, pp. 69-80.

Russell, J.S. and Moore, A.W., 1968. Comparison of different depth weightings in the numerical analysis of anisotropic soil profile data. Transactions of the 9th International Congress of Soil Science, 4: 205–213.

Sanchez, P.A., Ahamed, S., Carre, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonca-Santos, M.D., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vagen, T.G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A. and Zhang, G.L., 2009. Digital soil map of the world. Science, 325: 680-681.

Scull, P., Franklin, J., Chadwick, O.A. and McArthur, D., 2003. Predictive soil mapping: a review. Progress in Physical Geography, 27: 171-197.

Simonson, R.W., 1989. Historical highlights of soil survey and soil classification with emphasis on the United States, 1899-1970. International Soil Reference and Information Centre Technical Paper 18, Wageningen, The Netherlands.

Sirakov, N.M. and Muge, F.H., 2001. A system for reconstructing and visualising three-dimensional objects. Computers & Geosciences, 27: 59-69.

Skidmore, A.K., Ryan, P.J., Dawes, W., Short, D. and Oloughlin, E., 1991. Use of an expert system to map forest soils from a geographic information system. International Journal of Geographical Information Systems, 5: 431-445.

Smith, M.P., Zhu, A.X., Burt, J.E. and Stiles, C., 2006. The effects of DEM resolution and neighbourhood size on digital soil survey. Geoderma, 137: 58-69.

Solomatine, D.P. and Shrestha, D.L., 2009. A novel method to estimate model uncertainty using machine learning techniques. Water Resources Research, 45: Article Number: W00B11.

Stein, A. and Corsten, L.C.A., 1991. Universal kriging and co-kriging as a regression procedure. Biometrics, 47: 575-587.

Wackernagel, H., 1998. Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin.

Webster, R., 1977. Canonical correlation in pedology- how useful? Journal of Soil Science, 28: 196-221.

Webster, R., 2000. Is soil variation random? Geoderma, 97: 149–163.

Webster, R. and De La Cuanalo, H.E., 1975. Soil transect correlograms of North Oxfordshire and their interpretation. Journal of Soil Science, 26: 176-194.

Webster, R. and Oliver, M.A., 2001. Geostatistics for Environmental Scientists. John Wiley and Sons Ltd, West Sussex, England.

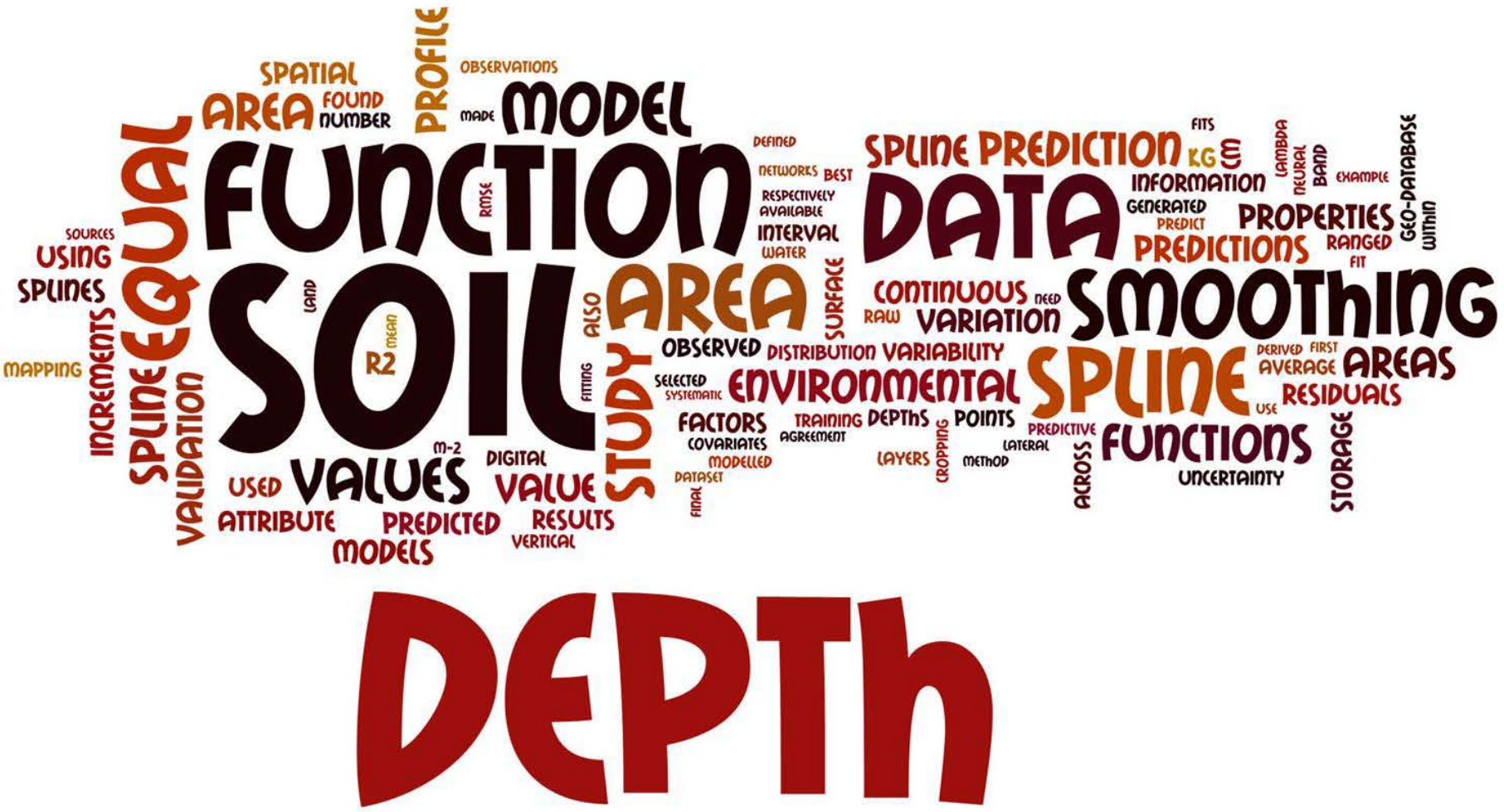
Webster, R. and Oliver, M.A., 2006. Modelling spatial variation of soil as random functions. In: S. Grunwald (Editor), Environmental Soil Landscape Modelling: Geographic Information Technologies and Pedometrics. Taylor and Francis, Boca Raton, pp. 241-288.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control, 8: 338-353.

Zhang, J. and Goodchild, M.F., 2002. Uncertainty in Geographical Information. Taylor and Francis, London.

Zhang, Y., Zhao, Y.C., Shi, X.Z., Lu, X.X., Yu, D.S., Wang, H.J., Sun, W.X. and Darilek, J.L., 2008. Variation of soil organic carbon estimates in mountain regions: A case study from Southwest China. Geoderma, 146: 449-456.

Zhu, A.X., Band, L., Vertessy, R. and Dutton, B., 1997. Derivation of soil properties using a soil land inference model (SoLIM). Soil Science Society of America Journal, 61: 523-53.


DEPTh
SOIL
FUNCTION
DATA
SMOOThING
SPLINE
AREA
MODEL
EQUAL
VALUES
ENVIRONMENTAL
FUNCTIONS
SPLINE PREDICTION
PREDICTIONS
PROPERTIES
INFORMATION
SPATIAL
AREA
FOUND
NUMBER
PROFILE
OBSERVATIONS
MADE
DEFINED
NETWORKS
BEST
RESPECTIVELY
AVAILABLE
INTERVAL
WATER
RMSE
SURFACE
ALSO
STUDY
FITTING
LAND
R2
MEAN
M-2
DIGITAL
VALUE
USED
ATTRIBUTE
PREDICTED
RESULTS
VERTICAL
MODELS
VALIDATION
INCREMENTS
SPLINES
USING
SOURCES
MAPPING
CONTINUOUS
NEED
RAW
VARIATION
OBSERVED
DISTRIBUTION
VARIABILITY
SELECTED
SYSTEMATIC
FACTORS
COVARIATES
MODELLED
DATASET
FINAL
TRAINING
AGREEMENT
DEPThS
POINTS
LAYERS
CROPPING
PREDICTIVE
LATERAL
METHOD
ACROSS
UNCERTAINTY
STORAGE
DERIVED
FIRST
AVERAGE
AREAS
RESIDUALS
USE
RANGED
FIT
GEO-DATABASE
WITHIN
EXAMPLE
LAMBDA
NEURAL
BAND
FITS
KG
CM
GENERATED
PREDICT

Soils have shape and area, breadth and width, as well as depth.
[Charles E. Kellogg 1949]

Chapter 2

Mapping continuous depth functions of soil carbon storage and available water capacity.

Summary

There is need for accurate, quantitative soil information for natural resource planning and management. This information shapes the way decisions are made as to how soil resources are assessed and managed. This chapter proposes a novel method for whole soil profile predictions (to 1m) across user-defined study areas where limited soil information exists. Using the Edgeroi district in north-western NSW as the test site, I combined equal-area spline depth functions with digital soil mapping techniques to predict the vertical and lateral variation of carbon storage and available water capacity (AWC) across the 1500km^2 area. Neural network models were constructed for both soil attributes to model their relationship with a suite of environmental factors derived from a digital elevation model, radiometric data and Landsat imagery. Subsequent fits of the models resulted in an R^2 of 0.44 for both carbon and AWC. For validation at selected model depths, R^2% values ranged between 20–27% for carbon prediction (RMSE: 0.30–0.52 log (kg/m^3)) and between 8–29% for AWC prediction (RMSE: 0.01m/m). In order to improve upon the model and validation results there is a need to address some of the structural and metrical uncertainties identified. Nevertheless, the resulting geo-database of quantitative soil information describing both lateral and vertical variation is an example of what can be generated with this proposed methodology. I also demonstrate the functionality of this geo-database in terms of data enquiry for user-defined queries.

2.1. Introduction

In order to benefit from the ecological and economical functions of soil in a sustainable way; land holders, private-sector and non-governmental stakeholders and, governmental departments need access to quantitative soil information. Such information confers weight to decisions regarding the management of the land and soil resources. Even in most environmental and agricultural research, accurate, continuous soil attribute data is becoming increasingly important in computer simulation models and for the assessment and monitoring of soil resources (Hempel et al. 2008). To facilitate this need, this chapter introduces a novel method for the prediction of user-defined, continuous soil properties to a given soil depth, across landscapes (at a fixed resolution) where only limited soil data exists.

The variation of soil properties down a profile is usually continuous (Ponce-Hernandez et al. 1986). Soil depth functions are often created to represent the depth-wise variation of soil properties. However, with traditional sampling of soil profile horizons, it is often assumed that the horizon value of a particular attribute represents the average value for that attribute for the depth interval of that horizon. With this paradigm, in effect what should be a continuous function, the data often appears discontinuous or stepped. Similarly, current digital soil mapping techniques are limited to map soil properties at specified depths or a combination of depth intervals (see Grimm et al. 2008 and Stoorvogel et al. 2009 as recent examples).

Bishop et al. (1999) pointed out that the discontinuity of depth functions derived from bulk horizon data leads to inaccuracies when attempting to predict the value of an attribute at specific depths within a soil profile. The earliest known attempt (to derive continuous depth functions of soil attributes) was by Jenny (1941) who drew freehand curves between attribute data points that corresponded to the mid-point value of a horizon. Over time, more sophisticated methods have evolved for constructing continuous soil depth functions such as using exponential decay functions (Russell and Moore 1968). Minasny et al. (2006) demonstrated that fitting exponential decay functions to carbon profile data resulted in an adequate quality of fit when attempting to map carbon storage in the Lower Namoi Valley, NSW. Linear regression and polynomials to the *n*-th degree by least-squares fitting have also been additional methods for deriving continuous soil depth functions (Colwell 1970; Moore et al. 1972). However, the disadvantage of these novel procedures is

that the value of a property at any depth affects the form of the fitted function at all depths (Ponce-Hernandez et al. 1986). As a consequence, the inflexibility of these functions results in a varied quality of fit (Webster 1978).

Irrespective of the soil property, a more flexible and accurate method for fitting continuous functions of soil data is the use of smoothing splines (Erh 1972) and equal-area spline functions as proposed by Ponce-Hernandez et al. (1986). Both Ponce-Hernandez et al. (1986) and Bishop et al. (1999) provide a good mathematical explanation of the operation of spline functions. Essentially, a spline function is a set of local polynomial functions- quadratic or cubic- tied together with 'knots' that describe a smooth curve through a set of points. Bishop et al. (1999) demonstrated their superiority over other continuous soil depth functions when they predicted a number of soil properties including soil pH, electrical conductivity (EC), clay content, organic carbon content, and gravimetric water content.

Clearly, continuous soil depth functions such as equal-area splines are advantageous for prediction of soil properties at specific depths. However, in a spatial context, a collection of spline functions for individual site observations will ultimately lead only to point observation data sets. To the parties concerned, such data will be of little use when they require continuous estimates of soil property variation across defined study areas or landscapes. The response to this demand has been answered partly in the way of digital soil mapping, where soil properties are mapped digitally based on their relationship with environmental variables (Minasny et al. 2008). The *scorpan* factors or environmental covariates as proposed by McBratney et al. (2003) provide a valuable predictive framework for determining soil variability in areas with limited soil data.

Given the predictive capabilities of soil depth functions and an explosion in the capabilities of digital soil mapping in areas with limited data (Lagacherie 2008), it seems only logical for there to be an amalgam of both methods to quantitatively predict the vertical and lateral variation of soil properties across a defined area. Using agronomically important soil properties-soil carbon storage and available water capacity (AWC) as the targets for prediction, this chapter proposes a novel method of predicting both their spatial and depth-wise variation. This is achieved in a number of stages:

- Fitting of equal-area spline functions to soil carbon and AWC profile data

- Assembly of a geo-database of environmental or *scorpan* factors for a defined study area where the point observations exist.
- Derivation of a neural network model using the best available set of *scorpan* factors to predict the depth-wise variation of the two soil properties.
- Extrapolation of the spline parameters onto the wider study area where soil observations do not exist.
- Mapping of the carbon storage and AWC of the entire study area to a depth of 1m.
- Demonstration of the functionality of the resulting soil geo-database for data enquiry.

2.2. Material and Methods

2.2.1. Study area

The study site is situated in the lower valley of the Namoi River, near Narrabri (30.32S 149.78E), approximately 500 km NNW of Sydney, NSW, Australia (Figure 2.2.1). Within this area, which covers approximately 1500 km^2, agriculture is the major landuse with irrigated cotton, wheat and pastoral farming being the predominant enterprises. Significant areas of native vegetation are also present, where it is mostly concentrated on the lower foothills of the Nandewar Range on the eastern flanks of the study site.

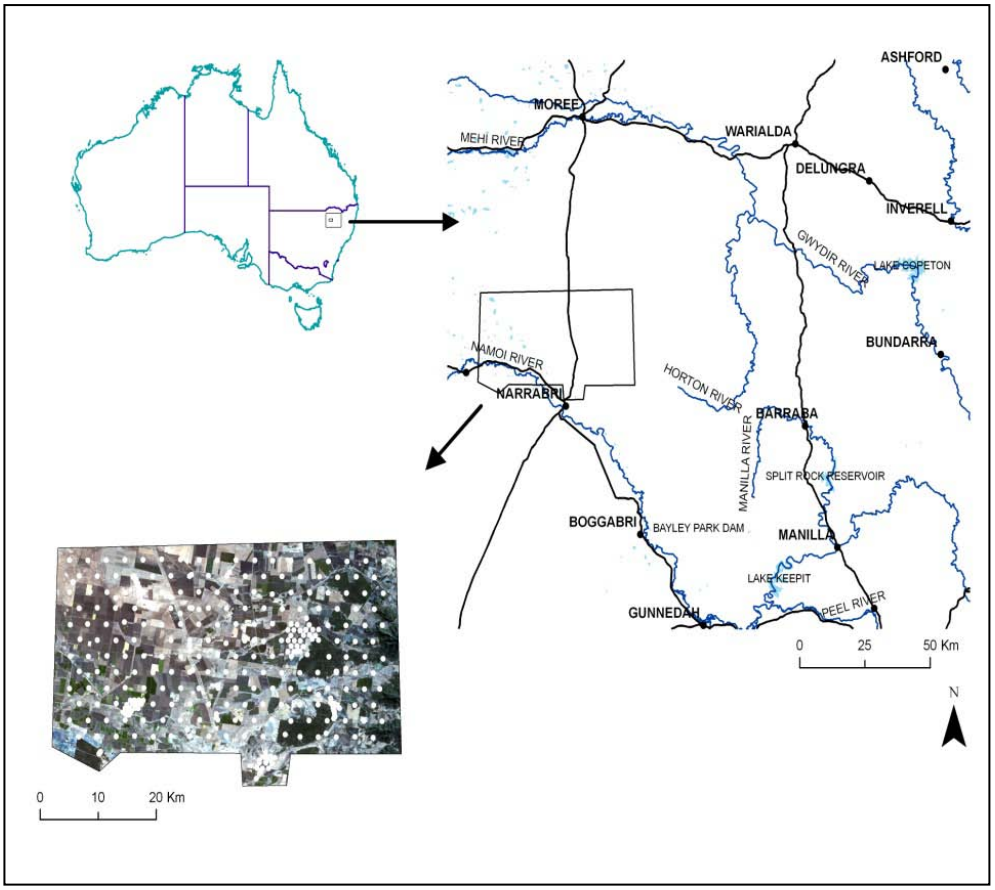


Figure 2.2.1. The Edgeroi study area. Inset image is a true colour satellite image with the sampling locations superimposed as white dots.

2.2.2. Environmental Data

For the purpose of digital soil mapping, a number of environmental indices were sourced and interpolated onto a common grid of 90 m resolution, encapsulated within the study area. These included:

- Landsat 7 ETM+ image taken towards the end of Summer 2003: The Enhanced Thematic Mapper Plus (ETM+) is a multispectral scanning radiometer that is carried on board the Landsat 7 satellite. There are 7 spectral bands that Landsat detects in the visible and near-infrared wavelengths: band 1 (0.45–0.52 μm), band 2 (0.52–0.60 μm), band 3 (0.63–0.69 μm), band 4 (0.78–0.90 μm), band 5 (1.55–1.75 μm), band 7 (2.09–2.35 μm). The Landsat bands were used to approximate the biosphere as a soil forming factor in terms of generalised land cover and land use. Vegetation cover and type was approximated using the Normalised Difference Vegetation Index (NDVI) determined by using

bands b3 and b4, where: NDVI = (b4-b3) / (b4+b3). Furthermore, the band ratios or more commonly, soil enhancement ratios of b3/b2, b3/b7 and b5/b7 were derived. It has been proposed that these soil enhancement ratios can accentuate carbonate radicals, ferrous iron, and hydroxyl radicals respectively in exposed soil (Saunders and Boettinger 2007).

- Gridded gamma-radiometric survey data were obtained from the recently compiled radiometric map of Australia, which integrates over 540 surveys at a 100 m resolution (Minty et al. 2009). The method of gamma-radiometric survey estimates the abundances of potassium (^{40}K), thorium (^{232}Th) and uranium (^{238}U) gamma-ray radiation emitted from the earth's surface. Cook et al. (1996) demonstrated that variations in the gamma-ray radiation of earth surfaces correspond with the distribution of various parent materials over the landscape.
- Digital elevation model (DEM) from the Shuttle Radar Topography Mission (SRTM) terrain data. From the DEM, first and second derivatives, namely: slope, aspect, terrain wetness index (TWI), flow length, slope length factor (LS-factor), area above channel network (AOCN) and stream power index (SPI) were determined. Moore *et al* (1993) and McKenzie and Ryan (1999) provide exemplary studies where some or all of the derivatives have been used to explore relationships with the spatial distribution of various soil properties.

2.2.3. The equal area smoothing spline

The spline model I used is a generalisation of the quadratic spline model of Bishop et al. (1999), where data are averages over adjacent horizons or layers in a soil profile. The model used in this chapter is generalised for cases where the supports of the data are not adjacent. Given measurements for soil properties at n layers in a soil profile, the boundaries of the layers are given in increments (u_1, v_1), (u_2, v_2), … (u_n, v_n), given that $u_1 < v_1 \leq u_2 < v_2 \leq \ldots \leq u_n < v_n$. The measurement of the bulk sample from layer i is assumed to reflect the mean attribute level, apart from measurement error. Mathematically, the measurements are modelled as:

$$y_i = \overline{f_i} + e_i \qquad \textbf{[2.2.1]}$$

where y_i is modelled estimate of the measured bulk sample from layer i. It is assumed that the true soil attribute values vary smoothly with depth. This is translated into mathematical terms. I denote depth by x, and the depth function describing the true attribute values by $f(x)$; which mean that $f(x)$ and its first derivative $f(x)$ are both continuous, and that $f'(x)$ is a quadratically integrable function. The depths of the boundaries of the n layers are given by $x_n < x_n, \ldots < x_n$, and $\bar{f}_i$ is the mean value of $f(x)$ over the interval (x_{i-1}, x_i) and e_i are measurement errors with mean 0 and variance σ^2. $f(x)$ represents a spline function, which can be found by minimising:

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{f}_i\right)^2 + \lambda \int_{x_0}^{x_n} [f_i(x)]^2 \, dx$$

[2.2.2]

The first term represents the fit to the data, the second term measures the roughness of function $f(x)$, expressed by its first derivative $f'(x)$. Parameter λ (lambda) controls the trade-off between the fit and the roughness penalty. The solution is a linear-quadratic smoothing spline, linear between layers, and quadratic within layers. See Appendix 2.1. for the derivation.

2.2.4. Mapping the smoothing spline soil depth function

The steps for mapping the smoothing spline soil depth function are as follows:

1) Collate legacy profile descriptions. They can have any kind of variation in any depth increments as long as they describe some of the variation in properties with depth. The soil dataset I used consists of 341 soil profiles. 210 were sampled on a systematic, equilateral triangular grid with a spacing of 2.8km between sites (McGarry et al. 1989). The further 131 soil profiles are distributed more irregularly or on transects (Figure 2.2.1). The dataset describes and quantifies various soil morphological, physical and chemical attributes at depth increments of 0–0.1, 0.1–0.2, 0.3–0.4, 0.7–0.8, 1.2–1.3 and 2.5–2.6 m.

   The focus of this study examines the vertical and lateral variability of carbon storage and AWC across the Edgeroi area at the time of measurement (1985-1987). The units of measurement used for carbon and AWC are kg m^{-3} and

m/m respectively for each depth interval. Soil organic carbon concentration was measured by McGarry et al. (1989). Using predicted soil bulk densities the soil organic carbon in kg m^{-3} was calculated. Soil bulk densities were estimated from a pedotransfer function using soil texture grades as predictors (Tranter et al. 2007). Similarly, AWC was derived from a pedotransfer function using sand, clay and organic carbon as predictors (Minasny et al. 2006).

2) Fit the spline to the values for each property. This generates a continuous profile description for each legacy soil profile with 1cm support. The maximum depth for the fitted spline was 1m. The spline function depends on a smoothing parameter- lambda (λ). Lambda values of 10, 1, 0.1, 0.01, 0.001, 0.0001 and 0.00001 were tried. The 'best' lambda value coincided with the spline that had the highest prediction quality i.e. the lowest root mean square error (RMSE). The 'best' lambda values were recorded and then assessed to determine the most frequently occurring λ for each variable. This value was then used as a blanket value to re-fit splines for all data points.
3) From the re-fitted spline, derive the mean value of the soil property within defined depth increments. For this study, the mean values at depth increments; 0–10, 10–20, 20–30, 30–40, 40–50, 50–70, 70–80 and 80–100cm were derived from the splines. The values for each property for each of the depth increments become the input for training data sets that are modelled with environmental covariates.
4) Implement a model framework to derive the relationship between the training data set and environmental covariates. Firstly, I joined the model inputs (based on their spatial location) to the environmental data using the nearest neighbour method. Stepwise regression was used to determine the best combination of environmental variables to predict both carbon content and AWC.

A regression kriging approach was used for this study. For the deterministic component of the kriging estimate, neural networks were used to model the training data as a function of the environmental covariates. Neural networks are a class of non-linear models that are composed of a large number of interconnected processing elements working in parallel to solve a particular

problem (Hastie et al. 2001). They are sometimes called single hidden layer back-propagation networks. The complexity of these models can be adjusted by the number of nodes within the hidden layer; the more nodes the more model parameters and subsequent model complexity. Successive neural networks were constructed for a training dataset for each soil attribute, each time adjusting model complexity by increasing the number of hidden nodes. I used neural networks on the basis that predictions can be made simultaneously at each depth interval for each iteration. For each iteration, 33% of the training data was used for cross-validation purposes in order to assess model over-fitting. Therefore, without compromising model predictive capability, I opted for models that also had a reasonable predictive capability for the cross-validated training data. The diagnostic of the model fits were reported as coefficient of determination or R^2. Table 2.2.1 below shows an example of this systematic model construction and selection process (using AWC as the example). Here I chose the third option (3 nodes) as with this configuration the greatest change in the R^2 value was achieved without greatly changing the cross-validated R^2.

Table 2.2.1. Examples of the accuracy of constructed neural networks for AWC with respect to iterative changes to the number of nodes.

Number of hidden nodes	R^2	Cross-validated R^2
1	9%	38%
2	34%	19%
3	44%	15%
4	52%	4%
5	59%	0%

Once selected, model formulae for each depth interval were saved for later use to predict in areas where data observations were not available. Regarding modelling the random or stochastic component for regression kriging, the residuals (i.e. the raw depth interval value derived from spline minus modelled value) were calculated. Ordinary kriging based on a global

variogram of the model residuals was then used to interpolate predicted residual values at each point on the common 90m grid.

5) Interpolate model rules or formulas onto the study area where information only relating to the environmental covariates exists. The kriged residuals (from step 4 of general procedure) were added to the predictions. Ultimately this resulted in a final prediction for each modelled depth interval (0–10, 10–20, 20–30, 30–40, 40–50, 50–70, 70–80 and 80–100cm) at each grid point within the study area. For each soil attribute, splines were reconstructed using the lambda parameter (from step 2) and predicted value of each soil attribute at each depth interval as inputs. From this, maps displaying the total cumulative value of each soil attribute were produced. Additionally, for demonstration of functionality, the resulting geo-database of soil information generated in this study was queried for the following three scenarios:

- At what depth does soil carbon concentration first decrease to below 1% (assuming a spatially constant bulk density of 1.3 g cm^{-3})?
- At what depth in the soil can we find the cumulative sum of carbon equal 5 kg m^{-2}?
- What is the lowest depth at which total AWC equals 0.1m?

  Maps were produced in order to visualise the results of each scenario.

2.2.5. Model validation

For model validation, the profile formulae were applied to 80 validation data points selected randomly from the original dataset. Residuals, coinciding with the location of each validation point were extracted from the 90m grid of residuals then added to the estimated depth values, equating to a final prediction. To visually assess the fit of predictions against observed legacy soil information, splines for selected validation data points were reconstructed using the 8 predicted depth increments and defined lambda parameter (from step 2) as inputs.

The diagnostic of validation is reported as the Root Mean Square Error (RMSE) which measures the differences between predicted and observed values. RMSE is equated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(z_{pi}(s) - z_i(s)\right)^2}{n}} \quad \textbf{[2.2.3]}$$

where $z_{pi}(s)$ and $z_i(s)$ are the predicted and true values of validation point *i* and *n* is the number of validation points, here 80. The other diagnostic used was Lin's concordance correlation coefficient (CCC) which measures the fidelity of the observations and the predictions to a 1:1 line (Lin 1989).

2.2.6. Implementation of methods

The fitting of splines to soil profile data, reconstruction of spline parameters for generating continuous soil profiles and running the user-defined queries were all implemented using **R**- a script based programming language (R Development Core Team 2011). JMP® 6.0.1 statistical software was used for the fitting of neural networks. Localised ordinary kriging was performed using the geostatistical software VESPER (Minasny et al. 2005). All other statistical procedures were implemented with **R**. All maps were generated using ESRI® ArcGIS software.

2.3. Results

2.3.1. Fitting equal-area splines to the dataset

The raw carbon data displayed a log-normal distribution and subsequently was log-transformed (natural log transformation) prior to fitting the splines. The data for AWC did not require any transformation. For carbon the standard deviation of the log-transformed data was found to be 0.89 with data values ranging from 0.0–4.7 log(kg m^{-3}). Overall, the best-fitting splines were found to have a lambda (λ) value of 0.01. AWC values ranged from 0.05–0.23 m m^{-1} with a standard deviation of 0.03. The best fitting splines for AWC were also found to have a λ value of 0.01.

2.3.2. Stepwise regression of environmental factors

Elevation, slope, radiometric K, band 5 and band ratios 3/7 and 5/7 were found to be strong covariates for both carbon and AWC prediction. Other environmental factors of importance for carbon prediction included altitude above channel network, stream power index and bands 3 and 4. While for AWC, LS-factor, terrain wetness index and NDVI were strong prediction covariates.

2.3.3. Neural networks for prediction of carbon and AWC from environmental factors.

From the systematic approach of neural network training, I selected models for carbon and AWC that featured 2 and 3 hidden nodes respectively as being the most appropriate (compromise between model predictive capability and over-fitting). Both neural networks resulted in fits where the R^2 value was 44%. Cross-validation of the training datasets resulted in R^2 values of 10% and 14% for carbon and AWC respectively. Depth wise, the prediction of carbon was best between 20–70 cm where the RMSE ranged between 0.27–0.30. The top of the soil profile was adequately predicted, with 50% of the variation for that layer explained. The predictions of the depth function for AWC were similar in that predictions were best between 20–70cm. The highest prediction errors were found at the top of the profile (0–10cm) followed by predictions at the bottom of the profile (80–100cm).

Model residuals at each depth interval were kriged using local neighbourhood prediction models. For both carbon and AWC, whilst there was some spatial patterning of residuals looking at each depth increment independently, there was no similarity in the spatial distribution of residuals when comparisons were made between each depth increment. Overall, a general observation was that there was only a slight degree of spatial auto-correlation of residuals for both soil attributes.

2.3.4. Model validation with the 80 withheld data points

Validation of AWC showed that model fits were significantly better in the top 20 cm of soil compared to the rest. The R^2 values for the top two depth increments prior to the addition of residuals were 28% (0–10 cm) and 25% (10–20 cm). For the remaining soil profile (20–100cm), R^2 values ranged between 6–12%. Figure 2.3.1a-c illustrate the observed vs. fitted plots at the selected depth increments of 0–10 cm, 30–40 cm and 80–100 cm respectively prior to the addition of residuals. At these depths, Lin's concordance coefficients range between 0.23–0.44 indicating a moderate agreement, with the strongest agreement for the 0–10cm depth increment. While there was some spatial pattern in the distribution of residuals (Figure 2.3.1d-f), their addition to predicted estimates of AWC made only a little improvement on the final predictions where R^2 values ranged between 8–29% (RMSE: 0.01). Lin's concordance coefficients (CCC) also indicate a modest improvement in predictions resulting from the addition of residuals (CCC: 0.27–0.49). Figure 2.3.1g-i illustrate

the observed vs. final prediction plots at the aforementioned selected depth increments.

Prior to the addition of residuals, validation of the neural networks on the 80 withheld data points indicate that the accuracy of carbon prediction decreased for each depth interval. For the top 40 cm of the soil profiles, model fits resulted in R^2 values between 17–24%. For the bottom 60 cm, R^2 values ranged between 13–15%. Despite the fact that there was not a very well defined spatial distribution of residuals at any significant separation distances > 160 m (data not shown), the addition of residuals to predictions had an overall improvement on model fits at all depth increments where R^2 values ranged between 20% and 27% (RMSE: 0.30–0.52).

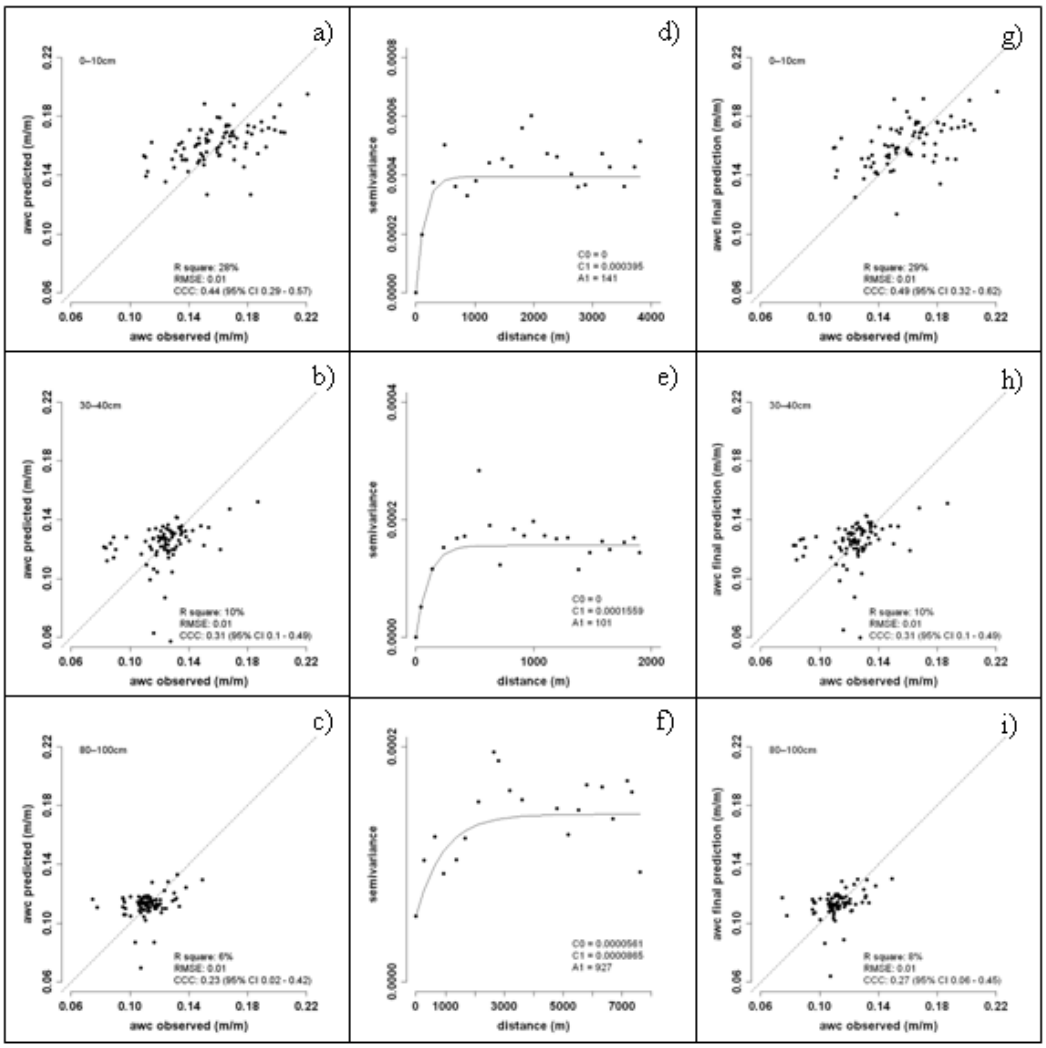


Figure 2.3.1. Available water capacity neural network prediction vs. observed plots from 80 randomly selected validation points at a) 0-10cm b) 30-40cm c) 80-100cm. Semi-variogram models of the residuals at each prediction depth (d-f). Final prediction (model prediction + residual) vs. observed plots at each prediction depth (g-i). RMSE: root mean square error. CCC: Lin's concordance correlation co-efficient.

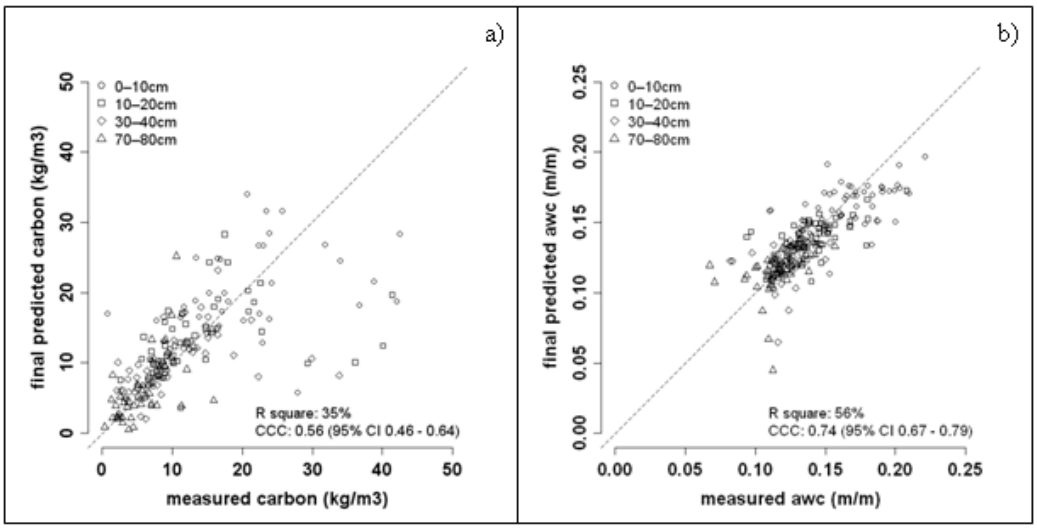


Figure 2.3.2. Validation of measured soil attribute (McGarry et al. 1989) vs. final predictions at 0–10cm, 10–20cm, 30–40cm, 70–80cm for a) carbon and b) AWC. CCC: Lin's concordance correlation co-efficient.

For an additional validation; the majority (not all) of the raw data of each soil variable were measured at specified depth ranges, for example, at 0–10cm, 10–20cm, 30–40cm and 70–80cm (McGarry et al. 1989). Validation results at these specific depth increments indicated good fits for both carbon and AWC (Figure 2.3.2). For carbon (R^2: 35%), the strongest agreement between the measured and final predicted carbon were in the 30–40cm and 70–80cm depth ranges. Conversely, for the 0–10cm and 10–20cm layers it can be seen that there were a greater proportion of systematic deviations from the 45° line. The resulting CCC of 0.56 is indicative of this. For AWC (R^2: 56%), there was overall a good agreement between the measured and final predicted values (CCC: 0.74), with no obvious deviations at specific depth increments.

Five data points were selected at random from the 80 validation points to graphically represent model predictions with observed data. These representations are shown in Figure 2.3.3f-j and Figure 2.3.4f-j for carbon and AWC respectively. The polygons represent the measured value at the specified depth increment. These figures are compared to the spline fits of the observed data calculated in the first stage of this study (Figure 2.3.3a-e and Figure 2.3.4a-e). Carbon values were back-transformed after the construction of the modelled spline estimates. Comparing the modelled spline functions fitted to the raw data indicate there is an average agreement between the predictions and the observed values. While not fitting

exactly to the raw data, the splines are sensitive to actual changes in carbon down the profile and follow the general trend of carbon distribution as from the raw data. For these graphical examples, the greatest errors of prediction were found at the surface, particularly at *ed002* and *ed044* (Figure 2.3.3f and g). At these sites the actual vs. predicted values of volumetric carbon differed by up to 7 kg m^{-3}.

For AWC prediction there was also only a fair agreement between the predicted spline depth functions and raw values (Figure 2.3.4f-j). In general the largest disagreement between predicted and actual values was found at the soil surface, particularly at *ed033, ed039, and ed157* (Figure 2.3.4f, g, i) where AWC differed by up to 0.05m/m.

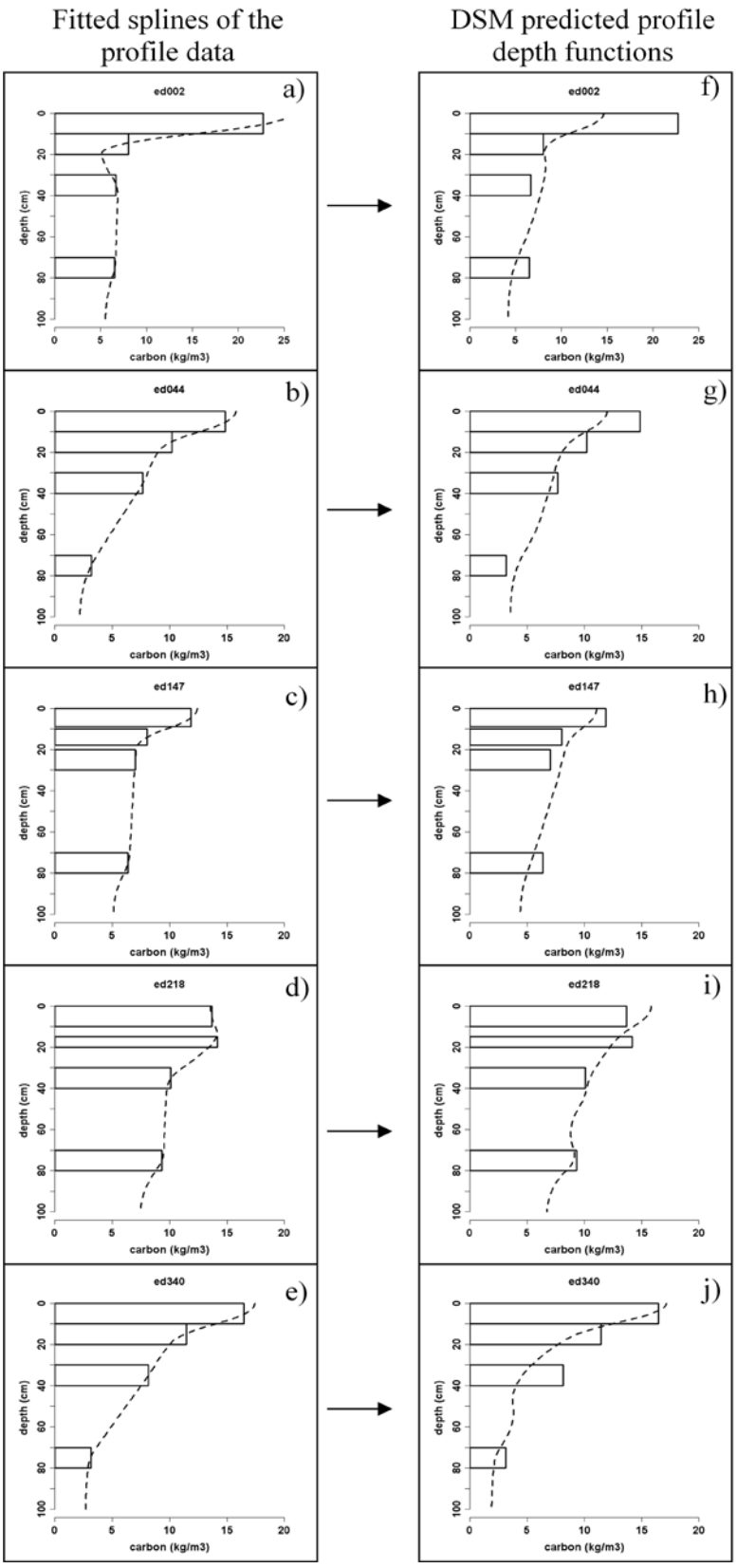


Figure 2.3.3. Fitted splines (dashed lines) of observed carbon profile data (polygons) at five randomly selected sites (a-e). Digital soil map prediction depth functions of carbon (dashed lines) and observed carbon profile data (polygons) at same selected sites (f-j).

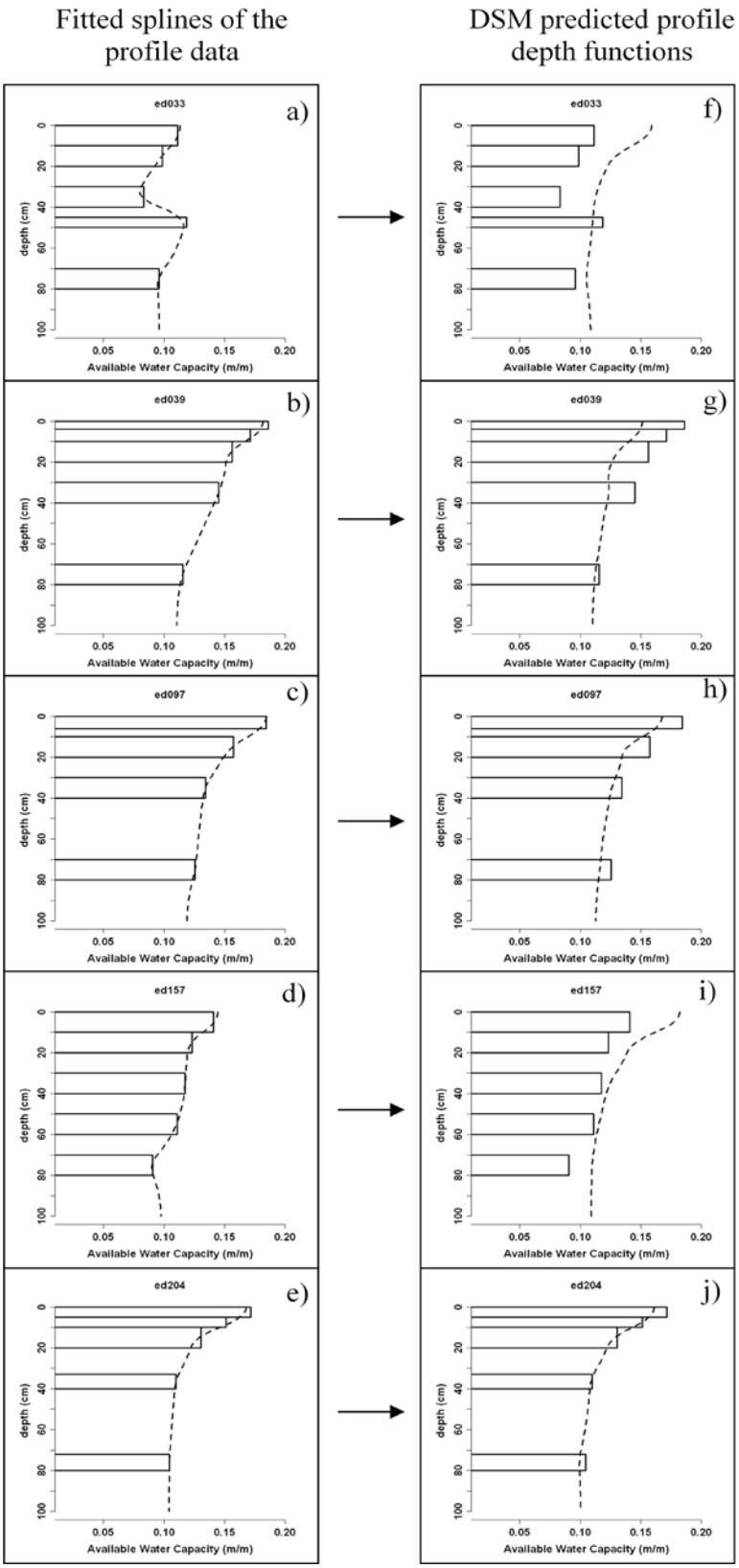


Figure 2.3.4. Fitted splines (dashed lines) of observed AWC profile data (polygons) at five randomly selected sites (a-e). Digital soil map prediction depth functions of AWC (dashed lines) and observed AWC profile data (polygons) at same selected sites (f-j).

2.3.5. Mapping carbon storage and available water capacity

Total carbon storage in the soil of the study area ranged between 1–50 kg m^{-2} to a depth of 1 m (Figure 2.3.5a). The total average carbon storage was 9.5 kg m^{-2}, with the highest levels found to the eastern and southern sections of the area (8–50 kg m^{-2}). From a generalised perspective these areas coincide with particular land uses not dedicated to cropping for example in forested areas, along watercourses and grazing areas. The cropping areas, situated in the northern and western sections of the area have the lowest carbon storage (1–7 kg m^{-2}). This trend is similar for the description of the spatial variability of AWC. Here, AWC ranged between 0.07–0.17 m to a depth of 1m, with an average of 0.13 m (Figure 2.3.5b). For the cropping areas to the north-west, AWC ranged mostly between 0.09–0.12 m.

The whole-soil profile maps (displayed as depth increments) of carbon storage and available water capacity are shown in Figure 2.3.6 and Figure 2.3.7 respectively. It can be observed that carbon is more abundant in the surface layers than in the sub-surface layers. The amount of carbon in the surface layer (0–10cm) of the cropping soils ranges between 6–9 kg m^{-3}. While for areas with natural vegetation, volumetric carbon ranges between 15–51 kg m^{-3} at the soil surface.

AWC in the 0–10 cm layer across the study area is relatively homogeneous where it ranges between 0.12 to 0.205 m m^{-1}, the lowest values occurring where cropping is practiced in the west of the study area. In all cases AWC decreases with the increase in soil depth. However, the rate of decrease in soil water is higher in areas where the land cover is either under native vegetation or pasture.

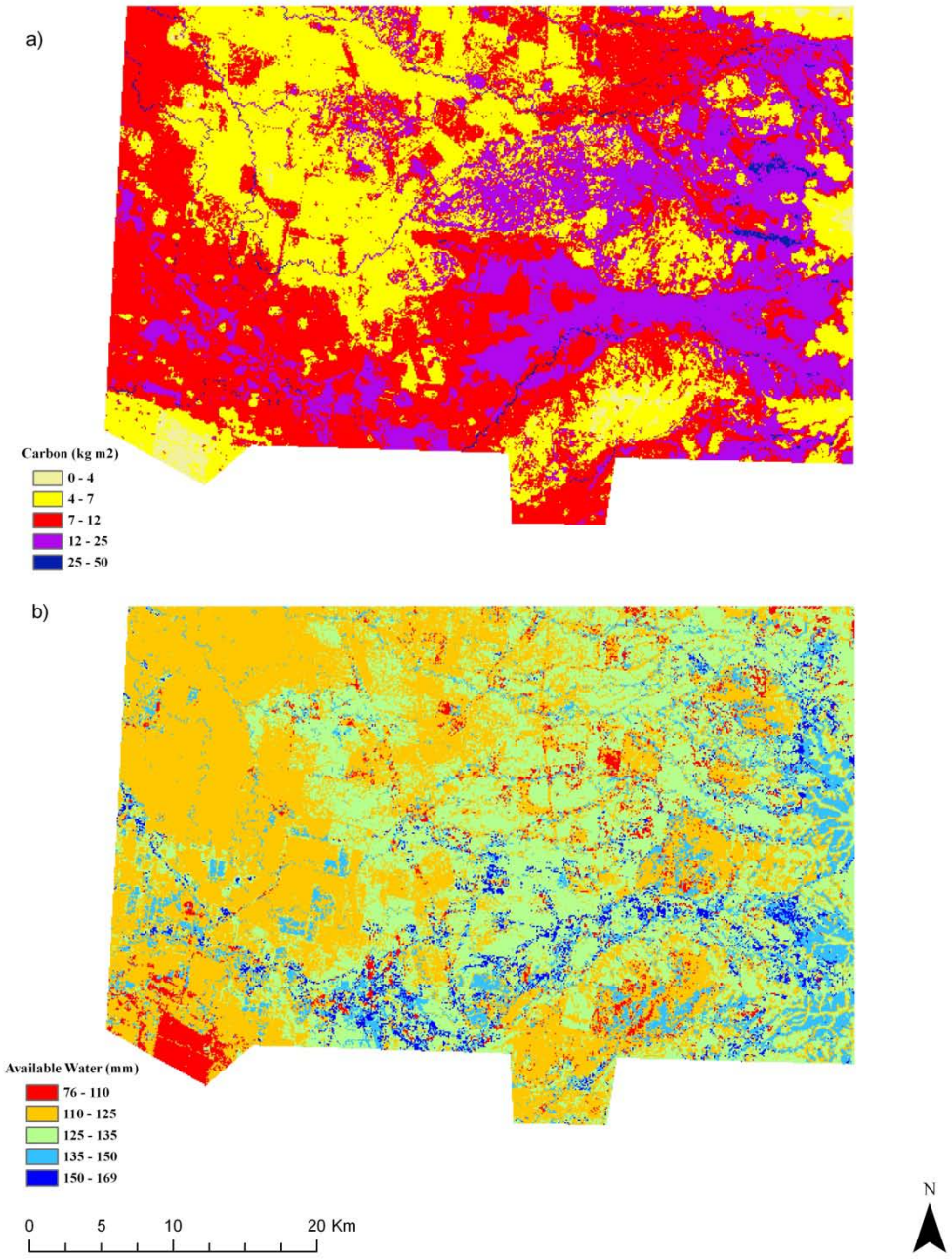


Figure 2.3.5. Predicted total carbon (kg m^{-2}) and available water capacity (mm) to a depth of 1m across the Edgeroi study area.

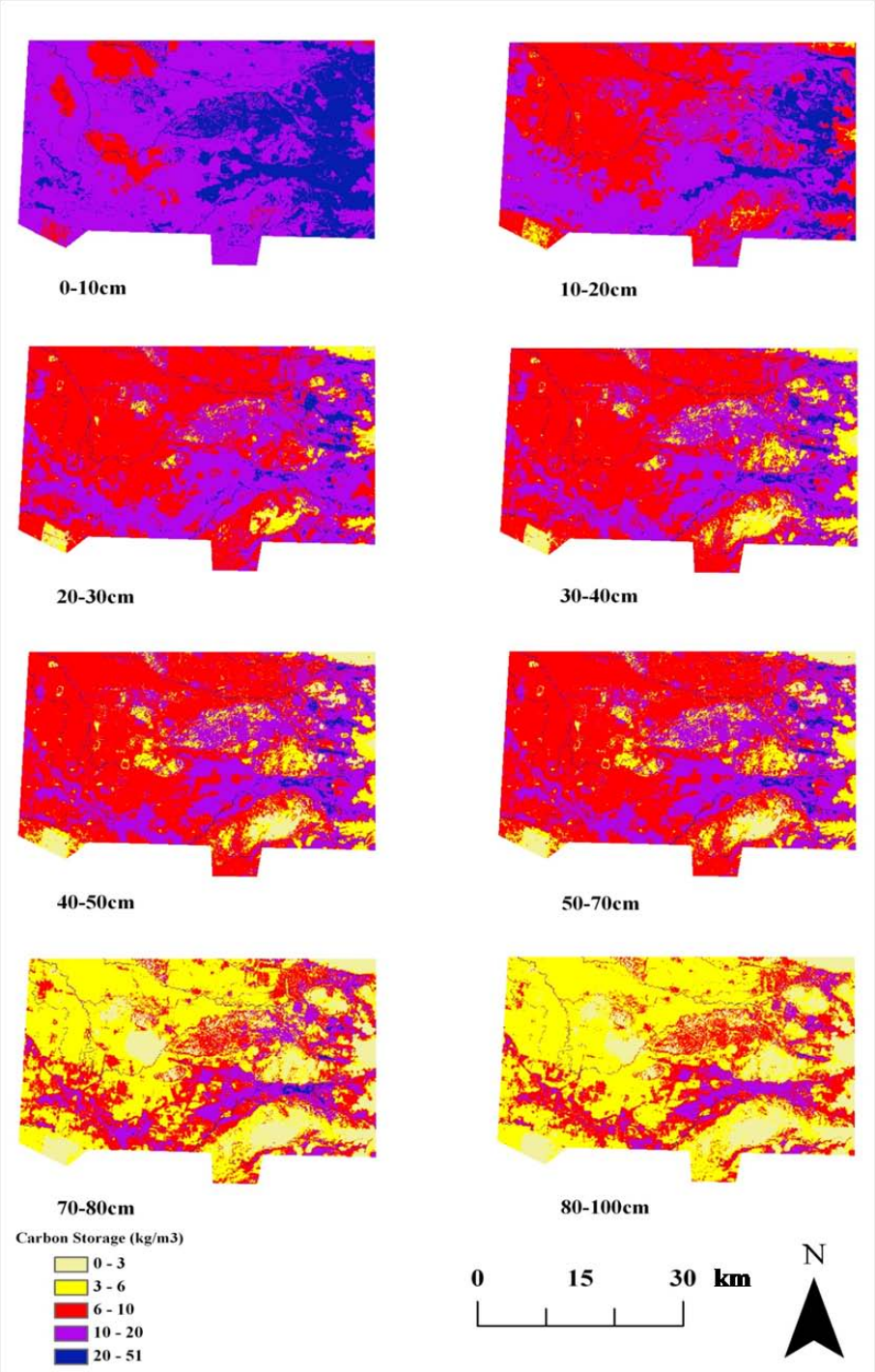


Figure 2.3.6. Predicted soil profile carbon (kg/m^3) to 1m displayed in 8 profile layers.

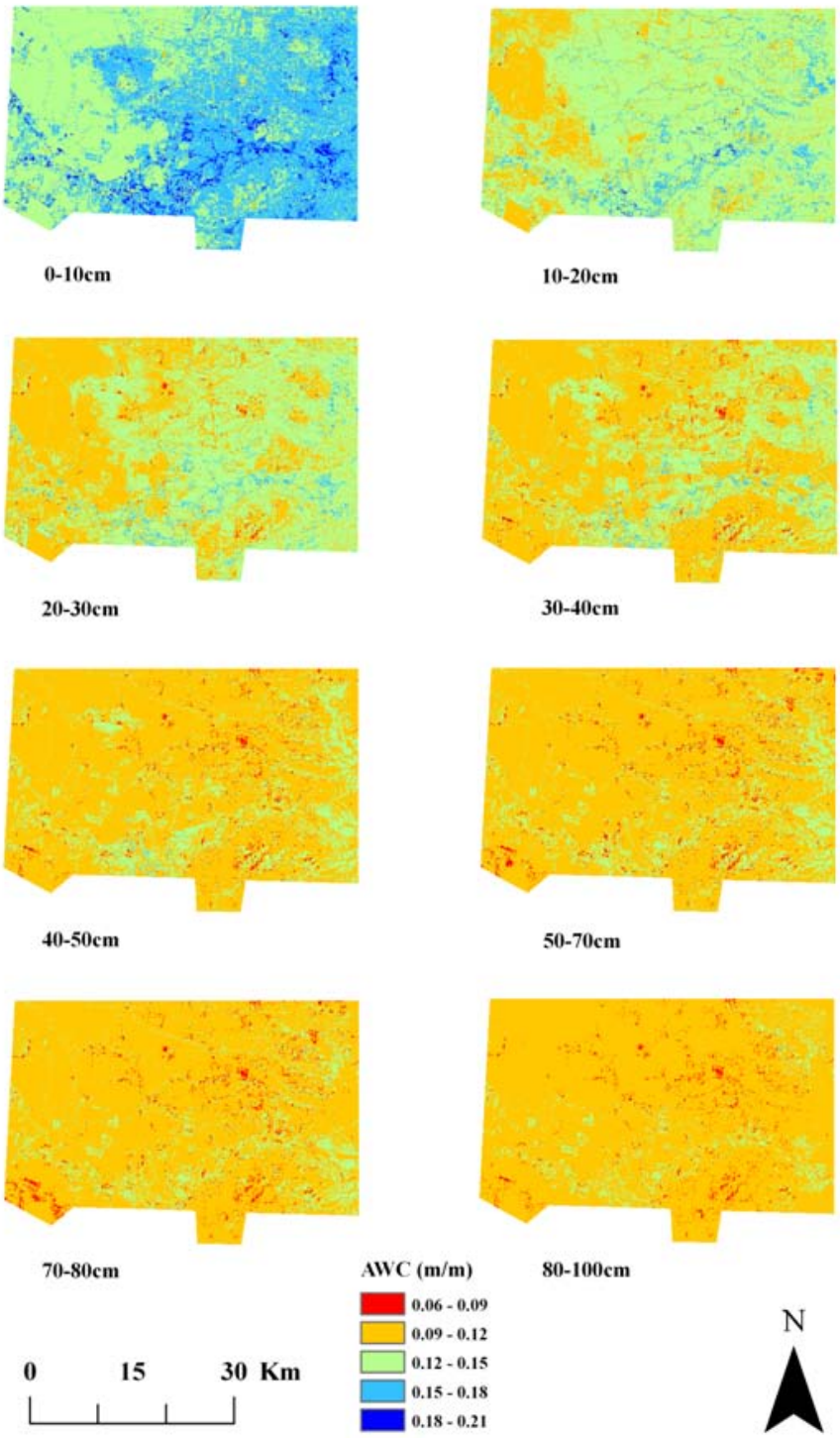


Figure 2.3.7. Predicted soil profile available water capacity (m/m) to 1m displayed in 8 profile layers.

2.3.6. Scenario-based queries of the generated soil geo-database.

By calculation, the geo-databases generated in this study each contain over 200 million pieces of information (210 370 lateral or grid square observations × 100 depth observations) that describe the vertical and lateral variability of carbon storage and AWC in the Edgeroi area.

As can be observed in Figure 2.3.8a. the depth at which soil carbon drops below 1% is quite variable across the study area, where it ranges from 1cm to over 1m, with the average depth at 21cm. The cropping areas situated mostly on the western areas of the study area tend to have the highest concentration of soils where in the top 5cm of soil, soil carbon falls below 1%. In most cases in these areas, even at 2cm below the surface, soil carbon has already decreased to below 1%. Conversely, the areas that do not appear to be cropped maintain soil carbon levels above 1% to greater depths. The range of depths at which soil carbon decreases to below 1% is much larger than that observed in the areas where cropping is practiced and would be predominantly due to land use (grazing as opposed to dense vegetation etc) and other factors such as parent materials and proximity to waterways.

There is some correlation between the depth at which soil carbon decreases to below 1% and depth at which the cumulative sum of carbon equals 5 kg m^{-2} (Figure 2.3.8b). This simple relationship highlights the negative exponential distribution of carbon in a soil profile, where carbon is most concentrated at the surface but decreases exponentially with increased soil depth. Therefore, the soils with low carbon storage i.e. greater depths required to attain 5 kg m^{-2}, are the ones that have initially low concentrations of carbon at the soil surface. Despite a number factors determining the variability of carbon across the study area the average depth required to attain 5 kg m^{-2} was found to be 50 cm.

In terms of the soil depth required to attain 100 mm of AWC (Figure 2.3.8c), the results indicated that the shallowest depth needed was 55 cm. However the frequency of this phenomenon occurring was relatively sparse and tended to be concentrated close to waterways or sources of water. Nevertheless the average depth required to attain 100 mm of AWC was 79 cm. There were some localised areas where greater than 1m was required. While variability in depth required does not appear significant, which could be due to a similarity in climate, a pattern of land use effect appears evident. Here the areas that have cropping have marginally less

AWC than those that do not, reflecting both an increased pressure on AWC to sustain crops and an increase in evaporation due to cultivation effects.

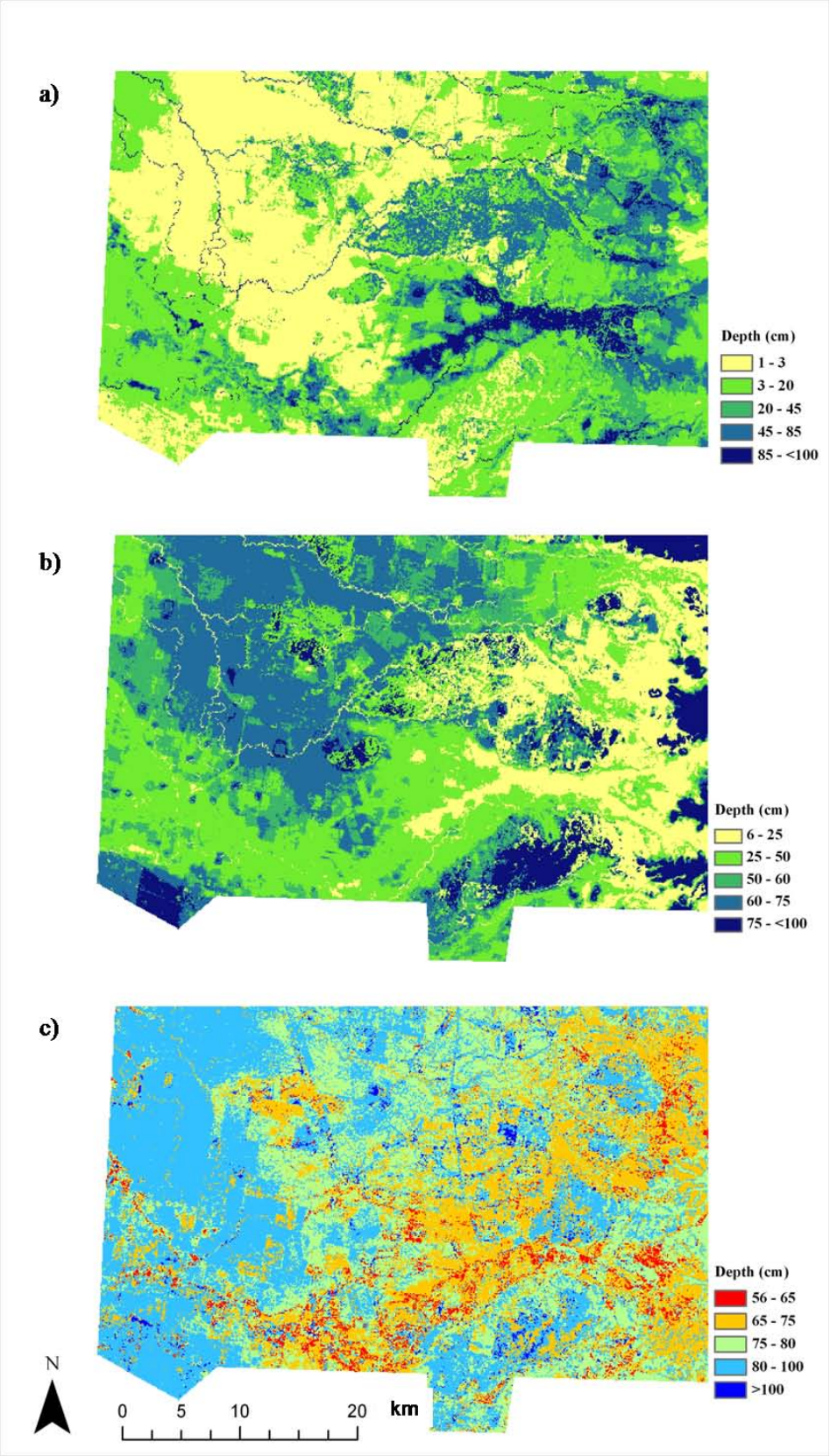


Figure 2.3.8. Maps of scenario-based queries. A) Depth at which soil carbon decreases to below 1%. B) Depth at which cumulative total of soil carbon equals 5kg m^{-2}. C) Depth at which cumulative sum of AWC equals 100mm.

2.4. Discussion

The Edgeroi area represents an ideal location for testing the methodologies of this study. First, the placement of survey sites in a mostly triangular grid means that site location is random with respect to the topography, landuse and soil type (McGarry et al. 1989). Secondly, the density of observations at each site ensures that the vertical distribution of soil properties is sufficiently represented.

Our prediction models of carbon and AWC were able to account for 44% of the variation of these properties across the study area. Similar accuracy assessments of models have also been reported for other digital soil mapping studies, for example, Ryan et al. (2000) and Florinsky et al. (2001). Broadly, these results are acceptable given that for quantitative soil spatial models, R^2 over 70% are unusual and values of 50% or less are common (Beckett and Webster, 1971). Our studies show that while the shape of the predicted spline depth functions is determined by the observed soil data, their flexible and sensitive nature makes them quite conducive for digital soil mapping purposes as we can fit them to any type of soil property and then derive information from them to use as model inputs. Post-modelling, the reconstruction of splines, results in densely populated datasets with fine vertical resolution (1 cm) in which can be queried as per the intentions of the user.

Validation results (R^2) of between 20-27% and 8–29% for carbon and AWC respectively indicated that the predictions were not as good as those generated by the prediction model. These results are confounding, given that our predictions at specific measured depths of the soil profile were of much greater agreement. Nevertheless, I did not expect to improve upon the results of the calibration models as we are trying to predict soil depth functions just from a suite of environmental variables. Many studies of mapping continuous soil properties are rarely validated. Studies that have, such as Minasny et al. (2006) (R^2= 50%), and Stoorvogel et al. (2008) (R^2= 8–23%) as examples, also reported slightly disappointing validation results. In this study, the slightly disappointing results could have stemmed from the fact that the models performed better for predicting surface soil conditions compared to sub-surface conditions (Minasny et al. 2006). Particularly for AWC, there was significant error propagation with depth. While the environmental variables that were used to predict carbon and AWC are important sources of variability, it is likely I have not fully realised other sources of variability. In other

words, I may have captured most of the variation of these soil properties at the soil surface using the existing data sources but I need to explore and seek out current and new data sources that will also explain their variation in the soil sub-surface. Additionally, I also found that the spatial distribution of residuals displayed very little meaningful variation. This affected the final model predictions in that only a slight improvement was achieved when they were incorporated into predictions. Other studies such as Odeh et al. (1995) and Stacey et al. (2006) have reported greater success using the regression kriging approach I implemented for this study.

The inability to capture completely, the variation of soil properties both the vertical and lateral space, should less be interpreted as a method failure, than as a limitation of the "low-cost" data input data set used in this study. It is common to use legacy soil data sets in DSM of the nature used in this study—often the quality is poorer in terms of sparseness and vertical resolution of observations. Thus one impetus for improving model performance and better detection of soil variability is to invest in further sampling from the mapping domain.

On the basis of this study, and the usage of low-cost data, one is still able to quantify the uncertainties or error in predictions relating to the soil spatial prediction functions. This uncertainty could be translated as being a combination of both metrical and structural uncertainty (Rowe 1994). Metrical uncertainty is unavoidable because of our reliance on models to define real objects. For example, the various terrain attributes calculated from a digital elevation model are based on a series of polynomial equations of various orders fitted using least-squares (Skidmore 1989). However, structural uncertainty in this study would be the most prevalent due to the fact that I am modelling environmental phenomena. Natural systems are inherently complex and difficult to define (Young 1998); a collection of environmental factors alone will never account for the entire variability of natural processes. In a lot of cases, models of natural systems do not account for interactions between factors, or more often there are a whole suite of other factors which are not considered or difficult to interpret (Young 1998). This brings us back to the point discussed previously where new and alternative data sources are needed if I want to capture more of the soil variability in both a lateral and vertical space.

While it is likely that improvements can be made if I follow these suggestions, what I really want is to be able to account for the known and unknown structural and metrical uncertainties within our predictions. Estimation of the uncertainty for

model outputs can be approached via a number of alternatives which include forecasting model outputs probabistically; analysing the statistical properties of model outputs and observed data; usage of simulation and re-sampling based techniques; or via methodologies based on fuzzy theory and machine learning techniques (Shrestha and Solomatine 2006). The study completed by Shrestha and Solomatine (2006) propose a novel method whereby they express uncertainty in the form of two quantiles (prediction interval). Chapter 3 will explore these ideas further, but expressing a prediction within a defined interval rather than having a single estimate seems an appropriate route to follow when we are modelling difficult environmental processes over large spatial scales.

Overall, the maps of carbon storage and AWC are the 'end-product' of a richly populated dataset of their variability in the Edgeroi area. At first glance, interpretations can be made to describe the pattern of variability. More importantly, behind these maps is an invaluable geo-database of quantitative soil information suited to the requirements of end-users for the assessment and monitoring of soil resources. The versatility of this data was demonstrated by the three scenarios that queried the underlying geo-database.

2.5. Conclusions

- Spline functions are sensitive and flexible to the variation of both carbon and AWC with soil depth and are thus quite amendable to use within the digital soil mapping framework.
- This study identifies two types of predictive uncertainty- structural and metrical. Our validation results indicate there is a need to address these forms of uncertainty. By incorporating a measure of uncertainty within predictions, improving the model calibration process and using new and existing alternative data sources as model variables, I envisage more reliable estimates can be generated to describe the lateral and vertical variation of soil in prescribed study areas.
- This study provides an example where a rich soil attribute geo-database can be generated from a limited soil dataset. I highlighted the functionality of this geo-database in terms of data enquiry for user-defined purposes.

2.6. References

Beckett, P.H.T. and Webster, R., 1971. Soil variability: a review. Soils and Fertilizers, 34: 1-15.

Bishop, T.F.A., McBratney, A.B. and Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, 91: 27-45.

Colwell, J.D., 1970. A statistical-chemical characterization of four great soil groups in southern New South Wales based on orthogonal polynomials. Australian Journal of Soil Research, 8: 221–238.

Cook, S.E., Corner, R.J., Groves, P.R. and Grealish, G.J., 1996. Use of airborne gamma radiometric data for soil mapping. Australian Journal of Soil Research, 34: 183-194.

Erh, K.T., 1972. Application of spline functions to soil science. Soil Science, 114: 333–338.

Florinsky, I.V., Eilers, R.G., Manning, G.R. and Fuller, L.G., 2002. Prediction of soil properties by digital terrain modelling. Environmental Modelling & Software, 17: 295-311.

Grimm, R., Behrens, T., Marker, M. and Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island - Digital soil mapping using Random Forests analysis. Geoderma, 146: 102–113.

Hastie, T.J., Tibshirani, R.J. and Friedman, J., 2001. The Elements of Statistical Learning: Data mining, Inference and Prediction. Springer, New York, NY.

Hempel, J.W., Hammer, R.D., Moore, A.C., Bell, J.C., Thompson, J.A. and Golden, M.L., 2008. Challenges to digital soil mapping. In: A.E. Hartemink, A.B. McBratney and M.L. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 81–90.

Ihaka, R. and Gentleman, R., 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics, 5: 299-314.

Jenny, H., 1941. Factors of Soil Formation: A System of Quantitative Pedology. McGraw-Hill, New York.

Lagacherie, P., 2008. Digital soil mapping: a state of the art. In: A.E. Hartemink, A.B. McBratney and M.L. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 3–14.

Lin, L.I.-K., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45: 255–268.

McBratney, A.B., Mendonca-Santos, M.L. and Minasny, B., 2003. On digital soil mapping. Geoderma, 117: 3-52.

McGarry, D., Ward, W.T. and McBratney, A.B., 1989. Soil studies in the Lower Namoi Valley: methods and data. The Edgeroi Dataset. CSIRO Division of Soils, Adelaide.

McKenzie, N.J. and Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma, 89: 67-94.

Minasny, B., McBratney, A.B. and Lark, M.R., 2008. Digital soil mapping technologies for countries with sparse data infrastructures. In: A.E. Hartemink, A.B. McBratney and M.L. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 15–30.

Minasny, B., McBratney, A.B., Mendonca-Santos, M.L., Odeh, I.O.A. and Guyon, B., 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. Australian Journal of Soil Research, 44: 233-244.

Minasny, B., McBratney, A.B., and Whelan, B.M., 2005. VESPER version 1.62. Australian Centre for Precision Agriculture, McMillan Building A05, The University of Sydney, NSW 2006. (http://www.usyd.edu.au/su/agric/acpa).

Minty, B., Franklin, R., Milligan, P., Richardson, M., Wilford, J., 2009. The Radiometric Map of Australia. Exploration Geophysics 40: 325-333.

Moore, A.W., Russell, J.S. and Ward, W.T., 1972. Numerical analysis of soils: a comparison of three soil profile models with field classification. Journal of Soil Science, 23: 193–209.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Science Society of America Journal, 57: 443-452.

Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes- heterotopic cokriging and regression kriging. Geoderma, 67: 215–226.

Ponce-Hernandez, R., Marriott, F.H.C. and Beckett, P.H.T., 1986. An improved method for reconstructing a soil-profile from analysis of a small number of samples. Journal of Soil Science, 37: 455-467.

R Development Core team, 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rowe, W.D., 1994. Understanding uncertainty. Risk Analysis, 14: 743-750.

Russell, J.S. and Moore, A.W., 1968. Comparison of different depth weightings in the numerical analysis of anisotropic soil profile data. Transactions of the 9th International Congress of Soil Science, 4: 205–213.

Ryan, P.J., McKenzie, N.J., O'Connell, D., Loughhead, A.N., Leppert, P.M., Jacquier, D. and Ashton, L., 2000. Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. Forest Ecology and Management, 138: 139-157.

Saunders, A.M. and Boettinger, J.L., 2007. Incorporating classification trees into a pedogenic understanding raster classification methodology, Green River Basin, Wyoming, USA. In: P. Lagacherie, A.B. McBratney and M. Voltz (Editors), Digital Soil Mapping: An introductory perspective. Developments in Soil Science. Elsevier, Amsterdam, pp. 389–399.

Shrestha, D.L. and Solomatine, D.P., 2006. Machine learning approaches for estimation of prediction interval for the model output. Neural Networks, 19: 225-235.

Skidmore, A.K., 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. International Journal of Geographical Information Systems, 3: 323–334.

Stacey, K.F., Lark, R.M., Whitmore, A.P. and Milne, A.E., 2006. Using a process model and regression kriging to improve predictions of nitrous oxide emissions from soil. Geoderma, 135: 107-117.

Stoorvogel, J.J., Kempen, B., Heuvelink, G.B.M. and de Bruin, S., 2009. Implementation and evaluation of existing knowledge for digital soil mapping in Senegal. Geoderma, 149: 161–170.

Tranter, G., Minasny, B., McBratney, A.B., Murphy, B., McKenzie, N.J., Grundy, M., Brough, D., 2007. Building and testing conceptual and empirical models for predicting soil bulk density. Soil Use and Management 23: 437-443.

Webster, R., 1978. Mathematical treatment of soil information. Transactions of the 11th International Congress of Soil Science, 3: 161–190.

Young, P., 1998. Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. Environmental Modelling & Software, 13: 105-122.

Appendix 2.1. Derivation of the equal-area spline function

The following is the derivation for the quadratic smoothing spline:
Given measurements for soil properties at *n* layers in a soil profile, the boundaries of the layers are given in increments (u_1, v_1), (u_2, v_2), … (u_n, v_n), given that $u_1 < v_1 \leq u_2 < v_2 \leq \ldots \leq u_n < v_n$. The measurement of the bulk sample from layer *i* is assumed to reflect the mean attribute level, apart from measurement error. Mathematically, the measurements are modelled as:

$$y_i = \bar{f}_i + e_i$$

[a2.1.1]

It is assumed that the true soil attribute values vary smoothly with depth. This is translated into mathematical terms. I denote depth by *x*, and the depth function describing the true attribute values by *f(x)*; which mean that *f(x)* and its first derivative *f'(x)* are both continuous, and that *f'(x)* is a quadratically integrable function.
The depths of the boundaries of the *n* layers are given by $x_n < x_n, \ldots < x_n$. Where $\bar{f}_i$ is the mean value of *f*(*x*) over the interval (x_{i-1}, x_i) and e_i are measurement errors with mean 0 and variance σ_2. The *f*(*x*) represents a spline function, which can be found by minimising:

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{f}_i\right)^2 + \lambda \int_{x_0}^{x_n} [f'(x)]^2 \, dx$$

[a2.1.2]

The quadratic spline

I define a quadratic spline *s*(*x*), in each layer; it conforms to a quadratic polynomial *p*(*x*). The polynomials $p_i(x)$ and $p_{i+1}(x)$ for two adjacent layers meet smoothly at the boundary. The curve is given by:

$s(x) = p_i(x)$ for $x_{i-1} \leq x \leq x_i$, $i = 1, 2, \ldots, n$.

[a2.1.3]

The smoothness conditions are:

$p_i(x_i) = p_{i+1}(x_i)$
$p_i^{'}(x_i) = p_{i+1}^{'}(x_i)$ for i = 1, 2, …, n – 1, and
$p_1(x_0) = 0$
$p_n^{'}(x_n) = 0$

[a2.1.4]

The latter two conditions mean that $s(x)$ is a natural spline. The points $(x_i, s(x_i))$ at the layer boundaries are called knots, and each x_i is referred to as a knot location. I define:

$f_i = s(x_i) = p_i(x_i)$
$b_i = s^{'}(x_i) = p_i^{'}(x_i)$
for
i = 1, 2, …, n.

[a2.1.5]

Quadratic polynomials

A quadratic polynomial can be written as:

$p(x) = \beta_0 + \beta_1 x + \beta_2 x^2$

[a2.1.6]

with coefficients $\beta_0, \beta_1, \beta_2$. In the case where the polynomial is over depth interval (u, v) where $u > v$, then the coefficients can be determined from:

$p^{'}(t),\ p^{'}(u),\ \overline{p} = \int_t^u p(x)dx/(u-v)$

[a2.1.7]

This is expressed as:

$p(x) = \overline{p} - \frac{p^{'}(u)+2\,p^{'}(v)}{6}\Delta + p^{'}(v)\,(x-v) + \frac{p^{'}(u)+2\,p^{'}(v)}{2\Delta}(x-v)^2$

for $v \le x \le u$, where $\Delta = (u-v)$.

[a2.1.8]

Smoothing quadratic splines

Since $f(x)$ is represented as a natural quadratic spline $s(x)$, $f'(x)$ is a linear function between knots. For any linear function $l(x)$, we have:

$$\int_u^t [f'(x)]^2\,dx = \frac{u-v}{3}(l(v)^2 + l(u)l(v) + l(u)^2)$$

[a2.1.9]

Hence:

$$\int_{x_0}^{x_n} [f'(x)]^2\,dx = \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{3}\left(b_{i-1}^n + b_{i-1}\,b_i \;\; + b_i^2\right)$$

[a2.1.10]

The condition that $f(x)$ is continuous at the internal knots yields:

$p_i(x_i) = p_{i+1}(x_i)$ for $i = 1, 2, \ldots, n-1$.

[a2.1.11]

Using (a2.1.8), this translates into a set of equations:

$b_{i-1}(x_i - x_{i-1}) + 2b_i(x_{i+1} - x_{i-1}) + b_{i+1}(x_{i+1} - x_i) = 6\left(\bar{f}_{i+1} - \bar{f}_i\right)$
for $i = 1, 2, \ldots, n-1$.

[a2.1.12]

This can be expressed in a matrix form. Let $\mathbf{R}$ be the $(n-1) \times (n-1)$ symmetric tridiagonal matrix with diagonal elements $R_{ii} = 2\,(x_{i+1} - x_{i-1})$ and off-diagonal elements $R_{i+1,i} = R_{i,i+1} = x_{i+1} - x_i$. Then:

$$\int_{x_0}^{x_n} [f'(x)]^2\,dx = \frac{1}{6}\mathbf{b}'\mathbf{R}\mathbf{b}$$

[a2.1.13]

Equation a2.1.2 becomes:

$$\frac{1}{n}\left(\mathbf{y} - \bar{\mathbf{f}}\right)^2 + \frac{\lambda}{6}\mathbf{b}'\mathbf{R}\mathbf{b}$$

[a2.1.14]

Minimising with respect to $\bar{f}$, the solution is represented as:

$$[\boldsymbol{I} + 6n\lambda(\mathbf{R}^{-1}\mathbf{Q}')'\mathbf{R}(\mathbf{R}^{-1}\mathbf{Q}')]\bar{\mathbf{f}} = \mathbf{y}$$

[a2.1.15]

Where $\boldsymbol{I}$ is the identity matrix, $\mathbf{Q}$ is a $(n-1) \times n$ matrix with $Q_{ii} = -1$, $Q_{i,i+1} = 1$ and $Q_{ij} = 0$ otherwise. Solving this equation yields the fitted layer values $\hat{\bar{f}}$. The fitted values at the knots can be obtained from:

$$\hat{\mathbf{b}} = 6\mathbf{R}^{-1}\mathbf{Q}'\hat{\bar{\mathbf{f}}}.$$

[a2.1.16]


PREDICTION
SOIL
MODEL
CLUSTER
DEPTH
UNCERTAINTY
CALIBRATION
INTERVAL
PI
OBSERVED
UNCERTAINTIES
EMPIRICAL
VALIDATION
LOWER
USED
OBSERVATIONS
LIMITS
INCREMENTS
PREDICTIONS
ENVIRONMENTAL
CONTINUOUS
APPROACH
COVARIATES
EXTRAGRADE
OC
MEMBERSHIP
PREDICTION
DISTRIBUTION
CLUSTERING
PREDICTED
PROCEDURE
UPPER
DATA
ALSO
DETERMINE
VALUE
INCREMENT
SPACE
OBSERVATION
ERROR
RESIDUALS
CLUSTERS
STUDY
CORRESPONDING
CONFIDENCE
SIMILAR
SPATIAL
FOUND
PERFORMED
DSM
FKM
MAPPING
OPTIMAL
PROFILE
PROCESS
USING
AREAS
INDEPENDENT
PIC ATTRIBUTES
ERRORS
CLASS
WITHIN
ALGORITHM
NUMBER
CENTROIDS
DATASET
FEATURE
METHODOLOGY
AVERAGE
VALUES
DEPTHS
DEFINED
MADE
BASED
METHOD
CHAPTER
LAND
BASIS
NEED
FUZZY
DERIVED
EXTRAGRADES
PL
PICP
AREA
FINAL
TERMS
AVAILABLE
ACROSS
FRAMEWORK
TRUE
GIVEN
SUITE
RESPECTIVELY
MEASURE
PROFILES

Since we can never know anything for sure, it is simply not worth searching for certainty; but it is well worth searching for truth; and we do this chiefly by searching for mistakes, so that we have to correct them.
[Karl Popper 1994]

Chapter 3

Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes.

Summary

In this chapter an empirical method is used where model output uncertainties are expressed as a prediction interval (PI) of the underlying distribution of prediction errors. This method obviates the need to identify and determine the contribution of each source of uncertainty to the overall prediction uncertainty. Conceptually, in the context of digital soil mapping, rather than a single point estimate at every prediction location, a PI, characterised by upper and lower prediction limits, encloses the prediction (which lies somewhere on the interval) and ideally the true but unknown value $100(1-\alpha)\%$ *of times on average the target variable (typically 95%). The idea is to partition the environmental covariate feature space into clusters which share similar attributes using fuzzy k-means with extragrades. Model error for predicting a target variable is then estimated from which cluster PIs are constructed on the basis of the empirical distribution of errors associated with the observations belonging to each cluster. PIs for each non-calibration observation are then formulated on the basis of the grade of membership each has to each cluster.*

I demonstrate how we can apply this method for mapping continuous soil depth functions. First, using soil depth functions and digital soil mapping (DSM) methods, I map the continuous vertical and lateral distribution of organic carbon (OC) and available water capacity (AWC) across the Edgeroi district in north-western NSW, Australia. From those predictions I define a continuous PI for each prediction node, generating upper and lower prediction limits of both attributes. From an external validation dataset, preliminary results are encouraging where 91% and 93% of the OC and AWC observations respectively fall within the bounds of their 95% PIs. Ideally, 95% of instances should fall within these bounds.

3.1. Introduction

From chapter 2, a novel method for digital soil mapping the whole-profile distribution of soil properties was introduced. This chapter extends this work by investigating how we might go about predicting the uncertainties of these predicted maps.

Soil scientists are acutely aware of the current issues concerning the natural environment because our expertise is intimately aligned with their understanding and alleviation. We know that sustainable soil management alleviates soil degradation, improves soil quality and will ultimately ensure food security. Critical to better soil management is information detailing the soil resource, its processes and its variation across landscapes. Consequently, under the broad umbrella of 'environmental monitoring', there has been a growing need to acquire quantitative soil information (McBratney et al. 2003; Grimm and Behrens 2010). The concerns of soil-related issues in reference to environmental management were raised by McBratney (1992) when stating that it is our duty as soil scientists, to ensure that the information we provide to the users of soil information is both accurate and precise, or at least of known accuracy and precision.

However, a difficulty we face is that soil can vary, seemingly erratically in the context of space and time (Webster 2000). Thus the conundrum in model-based predictions of soil phenomena is that models are not 'error free'. The unpredictability of soil variation combined with simplistic representations of complex soil processes inevitably leads to errors in model outputs.

We do not know the true character and processes of soils and our models are merely abstractions of these real processes. We know this; or in other words, in the absence of such confidence, we know we are uncertain about the true properties and processes that characterise soils (Brown and Heuvelink 2005). The key is therefore to determine to what extent our uncertainties are propagated in our models of these real-world processes.

In modelling exercises, uncertainty of the model output is the summation of the three main sources generally described as: model structure uncertainty, model parameter uncertainty and model input uncertainty (Minasny and McBratney 2002b; Brown and Heuvelink 2005). The general procedure is to determine independently the contribution of each source to the overall uncertainty. One obvious issue of this is

that generating estimates of uncertainty for each of the sources could become a prohibitive exercise in terms of time and cost.

With this in mind, there are a number of approaches to estimate the uncertainty of model outputs. One of these is an empirical approach in which the residuals between modelled outputs and corresponding observed data are used to formulate a prediction interval (PI). Such an approach was proposed by Shrestha and Solomatine (2006) where uncertainty is expressed in the form of two quantiles of the underlying distribution of model error (residuals). It is stated that the PI explicitly takes into account all sources of uncertainty and circumvents attempts to separate out the contribution of each source (Shrestha and Solomatine 2006; Solomatine and Shrestha, 2009). The purpose of the empirical methodology is to derive the upper and lower prediction limits based on the model errors, and since it is estimated through an empirical distribution, it is not necessary to make any assumption about residuals (Solomatine and Shrestha 2006). Their idea is to partition a feature space into clusters (with a fuzzy k-means routine) which share similar model errors. A PI is constructed for each cluster on the basis of the empirical distribution of residual observations that belong to each cluster. A PI is then formulated for each observation in the feature space according to the grade of their memberships to each cluster. They applied this methodology to artificial and real hydrological data sets and it was found to be superior to other methods which estimate a PI. The Shrestha and Solomatine (2006) approach computes the PI independently and while free of the prediction model structure, it requires only the model or prediction outputs. Tranter et al. (2010) extended this approach to deal with observations that are outside of the training domain.

Application of the Shrestha and Solomatine (2006) approach for estimating model output uncertainty has not previously been attempted in a digital soil mapping (DSM) framework where uncertainties are infrequently reported (Grunwald, 2009). Such was the case in Chapter 2 where a methodology for mapping continuous depth functions of soil attributes was introduced. Other than to identify the sources, no attempt was made to address the contribution of each source to the overall uncertainty. Given the modest results in terms of accuracy, particularly with increasing soil depth, it was assumed that significant model uncertainties existed. It is thought that, given the complexity of the modelling required to generate

predictions both in a vertical and lateral space, conventional forms of uncertainty analysis would be prohibitive and time consuming. A pragmatic approach to this dilemma is to use an empirical methodology in a similar fashion to that presented by Shrestha and Solomatine (2006) and Tranter et al. (2010) who view the model residuals as the best quantitative measure of the discrepancy between a model and the modelled real-world process. The efficiency of the method means it could be useful when we are dealing with soil spatial prediction functions that involve a regression kriging approach or machine-based models where there is no closed-form analytical solution for deriving model parameter uncertainties.

The method presented in this chapter modifies slightly the Shrestha and Solomatine (2006) and Tranter et al. (2010) approach to enable it for a DSM framework. It extends the idea of model uncertainty by extrapolating the uncertainty parameters across the extent of a defined area so that mapping the continuous depth functions involves both the mapping of predictions and their uncertainties. In addition to presenting this modified approach to uncertainty analysis, I also perform an external validation in order to gauge how successfully this method works for this particular application of DSM.

3.2. Theory and scope of work

3.2.1. The prediction interval as a measure of uncertainty

The characteristics of a PI include both upper and lower prediction limits. The interval between the prediction limits constitutes the PI (Figure 3.2.1). Given a prescribed probability such as a 95% confidence level, a future unknown value is expected to lie somewhere along this interval. There is a clear distinction between a PI and a confidence interval (CI), however. A CI tells us how well or how accurate is the estimate of a true regression to predict one variable from another. Conversely, the PI deals with the accuracy of the prediction with respect to the corresponding observed value. Thus a PI is always wider than a CI because it includes both the uncertainty in knowing the value of the population mean, and the uncertainty of the new measurement (Altman and Gardner 1988).

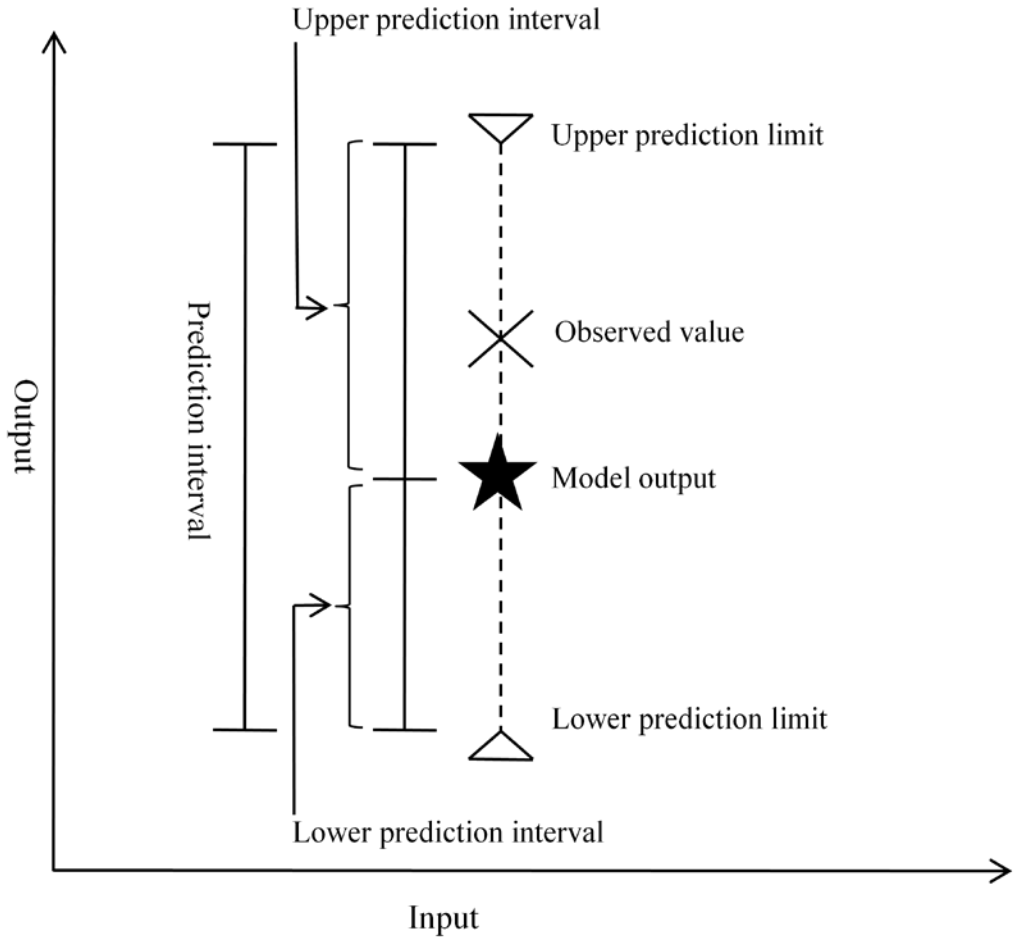


Figure 3.2.1. The prediction interval, characteristic features and general descriptive terminology. Adapted from Shrestha and Solomatine (2006).

Calculating the PI for a given observation following the method of Shrestha and Solomatine (2006) is performed independently of the model building or calibration process. Solomatine and Shrestha (2009) refer to this as ‘‘uncertainty estimation based on local errors and clustering’’ (UNEEC). The purpose of the UNEEC is to derive the upper and lower prediction limits based on the model error, and since it is estimated through an empirical distribution, it is not necessary to make any assumption about residuals (Solomatine and Shrestha 2009). First a user-defined class of regression is performed to estimate the target variable from one or a suite of predictor variables or covariates. The prediction outputs are compared to their observed corresponding values; the residuals are recorded. Using a clustering routine such as fuzzy k-means (Bezdek 1981), the calibration dataset is partitioned into *k* clusters corresponding to different values or distributions of the residuals. It is assumed that the region in the feature space associated with any particular cluster has

similar residuals or residuals with similar distributions. Once the clusters have been identified, the PIs for each cluster are computed from empirical distributions of the corresponding residuals. To construct a $100(1-\alpha)\%$ PI, the $(\alpha/2)\times 100$ and $(1-\alpha/2)\times 100$ percentile values are taken from the empirical distribution of residuals for the lower and upper prediction limits respectively. The computation of the PI for each calibration observation is straightforward if it belongs entirely to one cluster as would be the case where the input space is divided into crisp clusters e.g. hard clustering. However in the case of fuzzy clustering, where each observation belongs to all available clusters with respect to a membership grade, a "fuzzy committee" approach is used where the PI is computed using the weighted mean of the PI of each cluster (Shrestha and Solomatine 2006). This can be defined mathematically as:

$$PI_i^L = \sum_{i=1}^{c} m_{ij}\ PIC_j^L$$
$$PI_i^U = \sum_{i=1}^{c} m_{ij}\ PIC_j^{\mathrm{U}}$$

[3.2.1]

where PI_i^L and PI_i^U correspond to the weighted lower and upper PI for the *i*th observation. PIC_i^L and PIC_i^U are the lower and upper PIs for each cluster *j*, and m_{ij} is the membership grade of *i*th observation to cluster *j*. Finally, the lower and upper prediction limits (PL_i^L and PL_i^U respectively) are derived for each calibration observation by adding the prediction (from the prediction model) to PI_i^L and PI_i^U.

3.2.2. Validating the prediction interval

Validation of the PI is performed externally with a dataset separate from the calibration dataset. In the context of DSM, data splitting or collecting additional samples using some sort of probability sampling are the most common methods for which validations are then based upon; see Grinand *et al.* (2008) and Kempen *et al.* (2009) for recent examples of each. The procedure for validating (the PI) follows closely with Shrestha and Solomatine (2006) in which the idea is to simply determine whether each of the validating observed values is inside their respective prediction limits. By definition, the prediction limits enclose the true but unknown value $(1-\alpha)\%$ of times on average (typically 95%). The performance of the model is therefore evaluated by means of a prediction interval coverage probability (PICP)

(Shrestha and Solomatine 2006) whereby the PICP is the probability that all observed values fit within their prediction limits and is estimated by:

$$\text{PICP} = \frac{1}{V} \text{ count } i$$
$$i\text{: } PL_i^L \leq obs_i \leq PL_i^U$$

[3.2.2]

where V is the number of observations in the validation dataset. The clustering technique and uncertainty model is said to be optimal when the PICP value is close to $100(1-\alpha)\%$.

3.2.3. Fuzzy Clustering

Particularly important in this study is what methodology of clustering is used, especially in the context of soil variability and identifying regions in a study area where predictions are more certain in some areas in comparison to other areas. In general terms, clustering is the unsupervised partitioning of a feature space into natural groups or clusters which share some measure of similarity. Many clustering techniques are in existence, which Jain *et al.* (1999) comprehensively review. In the domain of soil science, the most widespread clustering algorithms are non-hierarchical or in other words, have a partitional basis (McBratney and Odeh 1997). The k-means algorithm is the simplest partitional clustering method and aims to minimise the within-class sum of square distances between the input space observations and the corresponding cluster centroids (McQueen 1967). An extension of the k-means algorithm is fuzzy k-means (FKM) which allows each observation a degree of membership to j clusters (Bezdek *et al.* 1984). Because soil is both spatially and temporally continuous, the FKM approach to classification intuitively appealing. The FKM algorithm minimises the objective function:

$$J(\mathbf{C}, \mathbf{M}) = \sum_{i=1}^{n} \sum_{j=1}^{c} m_{ij}^{\varphi} \, d_{ij}^2$$
$$i = 1, \ldots, n; \; j = 1, \ldots, c$$

[3.2.3]

where $\mathbf{C}$ is the $c \times p$ matrix of class centroids where c is the number of clusters and p is the number of variables; $\mathbf{M}$ is the $n \times c$ matrix of partial memberships, where n is the number of observations; $m_{ij} \in [0,1]$ is the partial membership of the ith observation to the jth cluster, $\varphi \geq 1$ is the fuzziness exponent. Increasing φ results in

a fuzzier partition between clusters. The square distance between the *i*th observation and *j*th cluster centre is d_{ij}^2. A more detailed explanation of the FKM algorithm can be found in Bezdek (1981). Essentially FKM gives the number of clusters; it defines class centroids based on each variable and calculates optimally the memberships of each observation to each defined cluster.

McBratney and de Gruijter (1992) recognised a limitation of the FKM algorithm in that it cannot distinguish between observations very far from the cluster centroids and those at the centre of the centroid configuration. The observations were termed extragrades as opposed to intragrades, which are the observations that lie between the main clusters. The extragrades are considered the outliers of the data set and have a distorting influence on the configuration of the main clusters (Lagacherie et al. 1997). McBratney and de Gruijter (1992) developed an adaptation to the FKM algorithm which distinguishes observations that should belong to an extragrade class. The FKM with extragrades algorithm minimises the objective function:

$$J_e(\mathbf{C}, \mathbf{M}) = \alpha \sum_{i=1}^{n} \sum_{j=1}^{c} m_{ij}^{\varphi} \, d_{ij}^2 + (1-\alpha) \sum_{i=1}^{n} m_{i^*}^{\varphi} \sum_{j=1}^{c} d_{ij}^{-2} \qquad \textbf{[3.2.4]}$$

The notation is similar to the FKM algorithm except where m_{i^*} denotes the membership to the extragrade class. This function also requires the parameter *alpha* (α) to be defined which determines the degree of importance attributed to the extragrade class. Details of FKM with extragrades are comprehensively discussed in McBratney and de Gruijter (1992) and Odeh *et al.* (1992).

Shrestha and Solomatine (2006) used a FKM algorithm for their clustering routine which also implemented the Euclidean distance measure where equal weight is given to all the variables in the feature space. For this study I implement the FKM with extragrades algorithm for the reasons described above. Tranter et al. (2010) also point out that extragrade instances exist spatially in regions of low density calibration data. As a consequence, this fact also confers a low reliability of prediction in these areas, which is an important consideration in the context of DSM. A Mahalanobis distance measure is also used in our procedure and for reasons discussed in more detail later is used on the basis that I cluster the feature space based on a suite of available soil prediction covariates rather than the prediction errors themselves. The

Mahalanobis distance takes into account the correlation between variables in the feature space.

3.2.4. Adaptation of the of the UNEEC approach for digital soil mapping of continuous depth functions

In order to use the UNEEC procedure of Shrestha and Solomatine (2006) and Solomatine and Shrestha (2009) within a DSM framework, some critical modifications and assumptions need to be made. The first involves the feature space clustering such that the model errors are calculated on the basis of the available soil state factors. The idea is to perform the clustering routine prior to running the prediction model after which the cluster PIs are then formulated. The key assumption of this chapter therefore is that particular areas within a landscape will have similar residuals or distribution of residuals and ultimately share a similar range of uncertainty.

The second modification or moulding to a DSM framework is the question of how I extend the PI to prediction nodes that have not been visited. Like in any DSM project, training rules are constructed on calibration data which are then extrapolated across a study area where only the prediction covariates are known. This approach is maintained for estimating the uncertainties at these sites whereby the cluster centroids derived from the calibration procedure are used to determine the membership grade of each prediction node across the study area to each cluster. Once these are known, PI_i^L and PI_i^U can be calculated after which PL_i^L and PL_i^U are derived once the model prediction is made.

In order to derive a continuous depth estimation of the upper and lower prediction limits for each prediction node, I follow the same routine as for mapping continuous functions of soil attributes as presented in chapter 2. Once the standard depths of prediction are determined, an uncertainty model is used to estimate the PI at each depth increment. Using the PIs at each depth as parameters, I can then perform a mass-preserving spline reconstruction to generate continuous representations of the upper and lower prediction limits to a prescribed maximum depth.

3.2.5. Procedure

There are thus three components that need to be adhered to in order to replicate this approach in a DSM framework. The first of these is the prediction model which

essentially recreates the method presented in chapter 2 for using splines and regression kriging modelling for the prediction of soil attributes at standardised depths. The second component involves the training of the empirical uncertainty model, from which cluster PIs can be derived. For both components, validation is performed using an independent dataset. In this study, both calibration and validation sets are the same for these two components. The third component involves the mapping of the predictions and associated PIs in the lateral and vertical dimensions. A summary of the stepwise procedure for achieving these outcomes is as follows and illustrated in Figure 3.2.2:

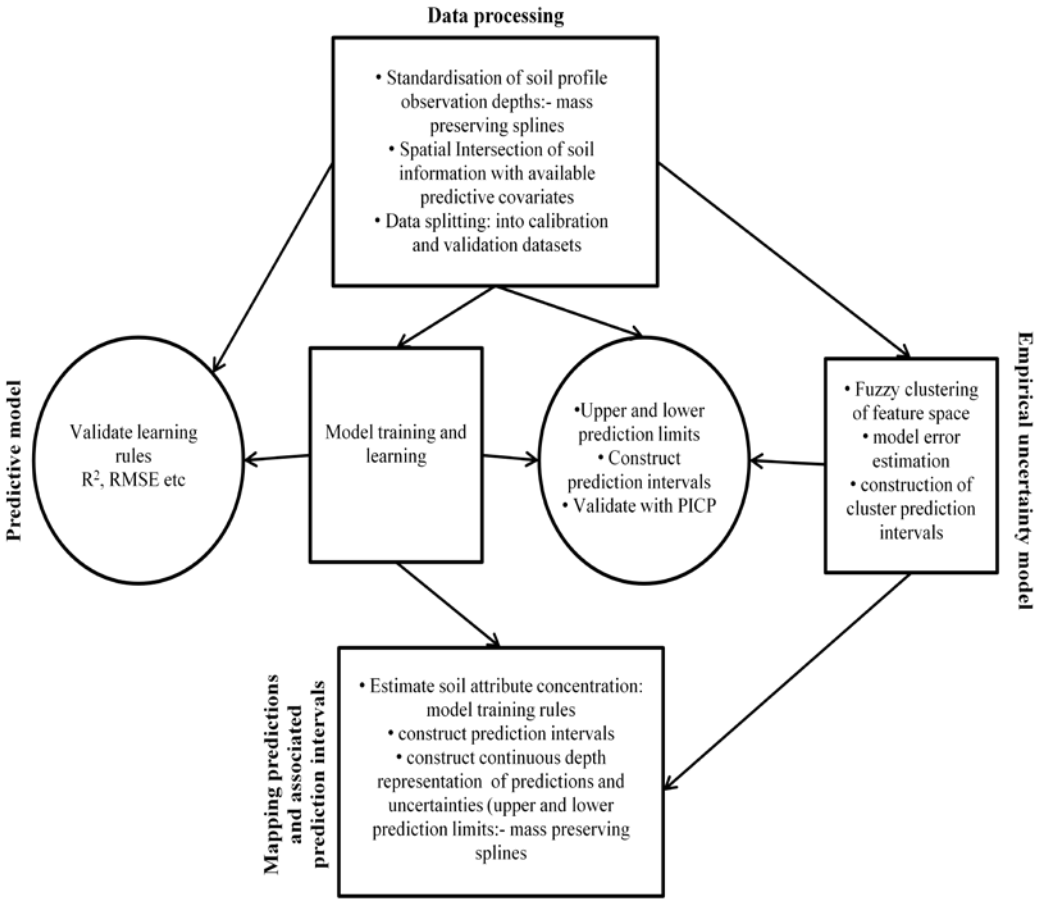


Figure 3.2.2. Flow diagram of the general procedure for achieving the outcome of mapping predictions and their uncertainties (upper and lower prediction limits) within a digital soil mapping framework. The 3 components for achieving this outcome are the prediction model, the empirical uncertainty model and the mapping component.

- Soil data is pre-processed then arranged for analysis. The framework I present here applies to available soil datasets where soil observations within a profile have

been made at horizon and/or regular depth increments. Often the depths of observation between profiles are not the same;

- Mass preserving spline functions (Bishop et al. 1999) are used to standardise the depth increments of prediction for all available site locations;
- The dataset is then spatially intersected with a suite of available environmental covariates;
- The dataset is then randomly split into two; I use 85% of total site observations for calibration and 15% for validation.

The prediction model

- The calibration dataset is used to train a prediction model via a regression technique in which the prediction variables are the environmental covariates;
- For validation of the prediction model, the training rules are extended to the validation data. Then evaluate statistically the accuracy of the predictions for each of the standard depths for example the root mean square error (RMSE; see equation 2.2.3 for derivation). It needs to be kept in mind that the validation data in this case are not the actual observations or 'hard data', but 'soft data' that in addition to sampling and measurement errors are prone also to interpolation errors due to the fitting of the splines. While not optimal, it is necessary to perform validation in this way (at the standardised depths) because the 'hard data' is rarely observed systematically at set depths from one profile to the next ;
- The training rules are further saved for the mapping component.

The empirical uncertainty model

- Clustering is performed where the feature space is the suite of environmental covariates observed at each calibration site. Once the optimal number of clusters is determined, cluster centroids are saved;
- Estimation of the model error is made. I would opt to use the same class of model as used in the prediction model component but generate predictions via a leave-one-cross-validation method. The idea is that we save only the residual between the predicted (model output) and observed value at each standardised depth;
- The highest membership value to a cluster determines which cluster each observation is assigned to. Cluster PIs (PIC_i^L and PIC_i^U) are found for each cluster

on the basis of the distribution of residuals within each cluster. To construct a 95% PIs (for each cluster) I take the lower 2.5% and upper 97.5% percentile values from the empirical distribution of residuals in each cluster;

- Validation (using the validation dataset) of the uncertainty procedure involves construction of the PI for each observation i.e. PI_i^L and PI_i^U (Equation 1) where the requirements are the cluster PIs and cluster centroids, from which the cluster membership values can be derived. Adding the prediction of the soil attribute (from prediction model) to PI_i^L and PI_i^U yields the upper and lower prediction limits (PL_i^L and PL_i^U) for each validation observation. Calculating the PICP determines (for a given confidence level) the proportion of observed values which fit within their respective prediction limits (Equation 3.2.2).

Mapping of the predictions and associated PI

- Training rules from the calibration procedure are extended to all prediction nodes across a study area where only covariate information is available;
- Cluster centroids from the uncertainty model are used to determine the membership grade each prediction node has to each cluster, from which PI_i^L and PI_i^U are formulated;
- Adding the model prediction to PI_i^L and PI_i^U yields the upper and lower prediction limits at each prediction node;
- Continuous depth representations of the predictions and upper and lower prediction limits can be generated by using the mass-preserving spline reconstruction method where the observations at the standardised depths are the only required parameters.

3.3. Material and Methods

3.3.1. The Data

I test the approach described above using actual soil data where our target properties are organic carbon (kg m^{-3}) (OC) and available water capacity (m m^{-1}) (AWC). The area from which the data has been collected is the Edgeroi district as described in section 2.2.1.

Environmental covariates were compiled for the whole Edgeroi study area (≈1500km^2) on a grid with spatial resolution of 90m × 90m. These included a digital

elevation model and its derivatives, Landsat 7 ETM+ images from 2003 and gamma radiometric data from airborne survey (Minty et al. 2009). Specifically, the environmental covariates used for analysis in this study were elevation, slope, altitude above channel network (AOCN), flow path length (FPL), multi-resolution index of valley bottom flatness (MRVBF) (Gallant and Dowling 2003) and SAGA wetness index which is similar to the topographic wetness index (TWI) (Boehner *et al.* 2002). Bands 1–5 and 7 of the Landsat 7 ETM+ were used in addition to the Normalised Difference Vegetation Index (NDVI) and soil enhancement ratios of b3/b2, b3/b7 and b5/b7 (Saunders and Boettinger 2007). For gamma radiometric data, the percentage measure of radiometric K was used in addition to the ppm measures of both radiometric U and Th.

For all soil profiles (341), the procedure as described in chapter 2 was used to fit splines to the observed values down to 1m. This generated continuous profile descriptions to 1m for both soil attributes. From the fitted splines of the observed data, the mean value of each soil attribute was derived at the specified depth increments of: 0–10, 10–20, 20–30, 30–40, 40–50, 50–70, 70–80, 80–100cm. These 8 surfaces became the target inputs to be modelled against the suite of environmental covariates which were then intersected to the data based on the spatial location of each soil profile description. Lastly, the dataset was then randomly divided into two sets: 291 profiles for calibration and 50 profiles for validation.

3.3.2. The prediction model

3.3.2.1. Prediction model calibration

In terms of the modelling process, the systematic approach for model calibration from chapter 2 was used to generate rules or formulae based on the relationship between calibration data at the eight specified depth increments and the suite of environmental covariates.

3.3.2.2. Prediction model validation

For validation, model formulae generated in the calibration procedure were used to derive the initial predictions of OC and AWC from the suite of environmental covariates that existed at each validating site. Kriging was used to interpolate the residuals at each validation observation based on the localised exponential variogram model of the 100 nearest residuals found for the calibration procedure. A final

prediction was derived from the summation of the model prediction and the interpolated residual. To determine the accuracy of the final predictions with their corresponding observed values I used the root mean square error (RMSE) and Lin's Concordance Correlation Coefficient (CCC) as described in section 2.2.5.

3.3.3. The empirical uncertainty model

3.3.3.1. Uncertainty model calibration: Fuzzy clustering and formulation of cluster prediction intervals.

To establish the optimal cluster size and φ value of the calibration data, fuzzy classification was performed with the *FuzME* software (Minasny and McBratney 2002a). As discussed previously, in this study I used the FKM with extragrades function. The environmental covariates of the calibration data were arranged in a matrix of *n* observations × *p* covariates (291×19). Iteratively, using cluster sizes of 2 through 15, the FKM with extragrades function ran using successive φ values ranging from 1 through 2 with step length of 0.01. For this study, along with an intuitive guide, I adopted an internal criterion approach to determine the optimal cluster size and φ value using both the modified partition entropy (MPE) and the derivative of the objective function with respect to the fuzzy exponent (φ) , $-(\delta J/\delta\varphi)c0.5$ (McBratney and Moore 1985). Such indices have been used previously by Odeh *et al.* (1992) and Bragato (2004) where more detailed discussion is made about them. The MPE establishes the degree of fuzziness created by a specified number of classes for a defined φ value. The notion is that the smaller the MPE, the more suitable is the corresponding number of classes at the given φ value. The derivative of $J_e(\mathbf{C},\mathbf{M})$ with respect to φ is used to simultaneously establish the optimal φ and cluster size. The optimal φ will maximise $-(\delta J/\delta\varphi)c^{0.5}$ and the most suitable cluster size will produce the curve with the lowest maximum. In this study these indices are used as a general guide; once the range of possible combinations has been narrowed, I then intuitively decided on the most suitable cluster size and φ value on other non-clustering related criteria such as the number of observations within each cluster and the associated distribution of errors assigned to each cluster i.e. as close to being normally distributed as manageable.

To determine model error I employed a leave-one-out-cross-validation (LOCV) procedure (Hastie *et al.* 2009). With this form of cross-validation, there were *n*=291

sets of computations. Using neural networks, a prediction is made for each successive calibration profile (at each depth increment) based on the learning rules of the remaining *n-1* calibration profiles. Furthermore, with each computation, kriging was used to interpolate the residual at each depth increment based on the spatial auto-correlation of residuals of the *n-1* calibration profiles. A final prediction resulted from the summation of the prediction and interpolated residual. Thus model error in this case was determined to be the difference between the observed value at a specific depth increment and its corresponding final prediction.

To calculate the cluster PIs (PIC_i^L and PIC_i^U) at each depth increment, I first arranged the observations into their respective clusters on the basis of their highest cluster membership grade. For a 95% PI, I took the upper 97.5% and lower 2.5% quantiles of each distribution for every cluster. In terms of handling the extragrade cluster error distributions, I follow the procedure of Tranter et al. (2010) whom suggested a penalisation calculation to extragrade areas where there is very low model prediction confidence. The PI for the extragrade class can be evaluated by:

$$PIC_{ej}^L = 2 \times q_{2.5}$$
$$PIC_{ej}^U = 2 \times q_{97.5}$$

[3.3.1]

where PIC_{ej}^L and PIC_{ej}^U are the lower and upper PIs of the extragrade class, and q is the quantile value of the extragrade cluster error distribution at each depth increment.

3.3.3.2. Uncertainty model validation

Validation of the uncertainty model follows the description as detailed in section 3.2.5. where $PI_i^L$and PI_i^U are formulated. The PICP is estimated accordingly for a 95% PI based on the count of observed values that lie within the PI for each site at each depth. As such the PICP considers all observations (site and depth increments) as independent observations. Thus the PICP is the proportion at all depths across all observations which lie within the 95% PI. To assess the sensitivity of the model by means of reducing the confidence limit sequentially, I constructed PIs for various confidence levels ranging from 5% to 99%. As for a 95% prediction level, by definition you would expect the PICP value to be close to the corresponding confidence or $100(1-\alpha)\%$ level.

3.3.4. Mapping of predictions and their uncertainties

Mapping of the predictions and associated prediction PIs follows precisely the steps as outlined in section 3.2.5. Along with generating maps of the predictions and their uncertainties (displayed as upper and lower prediction limits) at the standard depths I also demonstrate the functionality of the splines in a DSM framework which was also demonstrated in chapter 2. In this study I determine total predicted AWC and OC across the study area to a depth of 1m. These predictions are also accompanied by upper and lower prediction limits.

3.3.5. Implementation of methods

All statistical methods carried out in this study were performed with either **R** or Matlab®- a computer programming language software. Matlab® was used specifically for the neural networks and thus the LOCV procedures. FuzME (Minasny and McBratney 2002a) was used to perform fuzzy k-means with extragrades. All maps were generated using ESRI® ArcGIS software.

3.4. Results and Discussion

3.4.1. The prediction model

3.4.1.1. Prediction model calibration

For model calibration of OC, I found that a neural network model with 4 hidden nodes was appropriate in terms of predictive power without over-fitting the data. For AWC, 3 nodes was found to be the most appropriate model configuration. The coefficients of determination (R^2) were reasonable at 54% for OC and 48% for AWC.

3.4.1.2. Prediction model validation

Generally predictions of OC were strongest at the surface and poorest at the bottom of the soil profile. As shown in Figure 3.4.1a-c, CCC ranged from 0.28 (RMSE=0.3) in the 0–10cm depth increment, to 0.11 (RMSE = 0.54) at 30–40cm through to -0.05 (RMSE = 0.90) for the 80–100cm depth increment.

It was apparent that there was a low spatial auto-correlation between residuals at all 8 depth increments (data not shown), which was also observed in chapter 2. Nevertheless, adding the model prediction to the interpolated validation residuals

resulted in modest improvements to the prediction of OC. In Figure 3.4.1d-f, CCC increased to 0.38 and 0.14 at the 0–10cm and 30–40cm depth increments respectively, but no change was observed at 80–100cm. In each of the cases there was no improvement in the RMSE values.

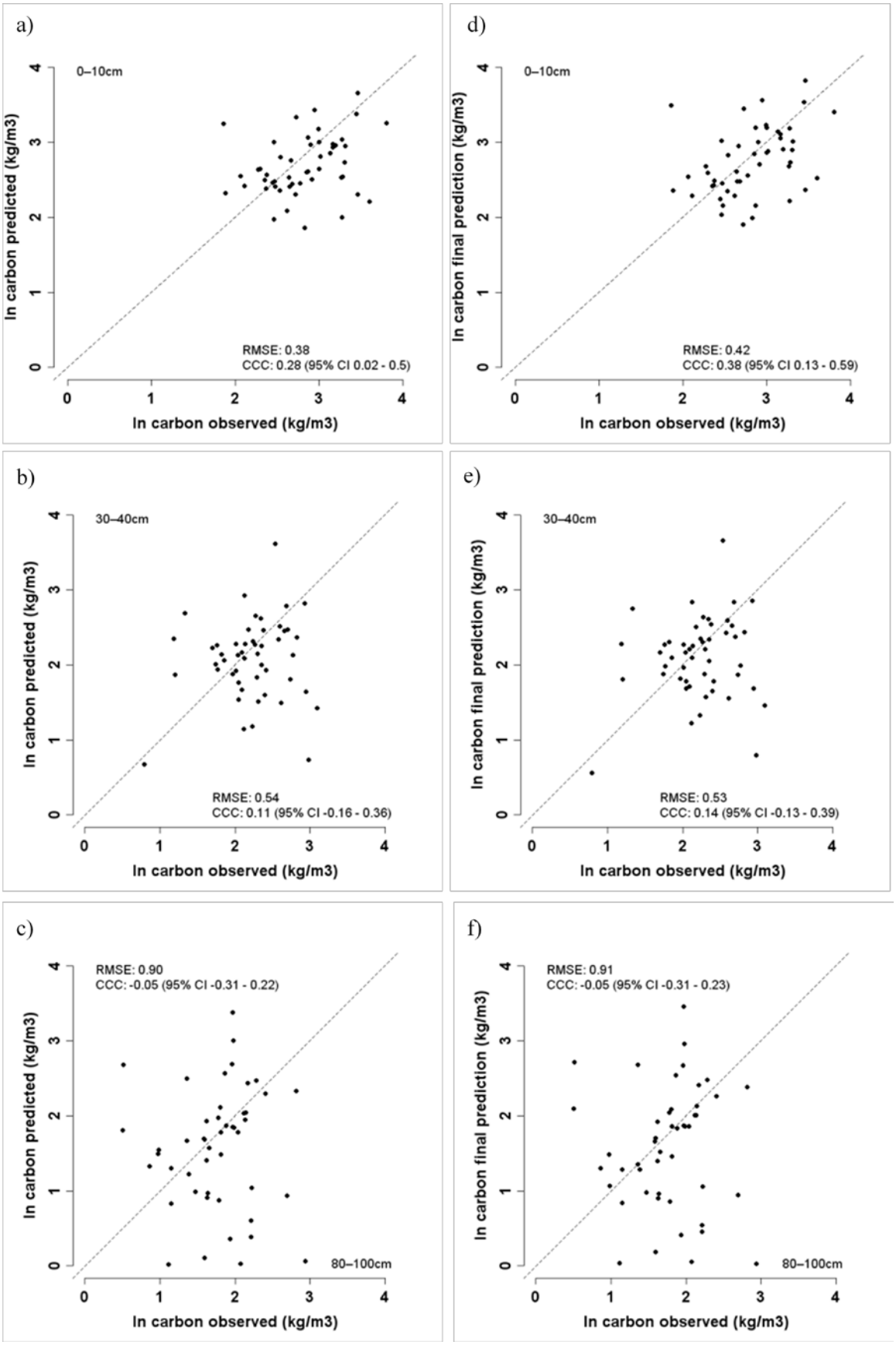


Figure 3.4.1. Model validation of OC. Observed vs. fitted plots at 0–10cm, 30–40cm and 80–100cm before adding residuals (a-c) and after adding residuals (d-f).

For AWC, the validation results were also modest where predictions were strongest in the 80–100cm depth increment (CCC=0.1) (data not shown). In terms of all 8 depth increments CCC ranged between 0.04–0.10 (RMSE = 0.01–0.03). As for OC, there was not a well defined function for the spatial distribution of residuals. Additionally, any spatial auto-correlation that I was able to define was independent for each depth increment. Overall, adding the residuals to the predictions of AWC resulted in little to no improvement.

The average CCC of the final validation predictions at the 8 depth increments for AWC and OC found in chapter 2 was 0.44 and 0.38 respectively. Essentially the validation results in this chapter are weaker. An explanation for this is that I did not perform independent multivariate analyses for both attributes to determine what the most correlated covariates were before construction of the neural networks. This is because, as discussed later, I performed the clustering process of the environmental covariates once only. Thus the defined clusters could be used simultaneously for both soil attributes. Given that the final predictions of both OC and AWC are modest in this chapter, future studies of this type will need to include an independent multivariate analysis prior to modelling. This would also mean that the clustering process would become independent for each predicted soil attribute and that the error determination through the leave-one-out-cross-validation would only include those environmental covariates that are significant for each soil attribute.

3.4.2. The empirical uncertainty model

3.4.2.1. Uncertainty model calibration: Fuzzy clustering and formulation of cluster prediction intervals.

The optimal cluster size for the given environmental covariates using the FKM with extragrades algorithm was found to be 6, including the extragrade class. Clustering resulted in 61 observations belonging to cluster A, while 38, 43, 54 and 61 observations belonged to cluster B, cluster C, cluster D, cluster E and the extragrade cluster respectively.

Box plots of the empirical distribution are shown for AWC (Figure 3.4.2a-c) and OC (Figure 3.4.2d-f) at the depth increments of 0-10cm, 40-50cm, and 80-100. At all depth increments the distribution of model errors are different for each cluster. For both soil attributes there was a decreasing distribution of model errors with

increasing depth down the profile, which was proportional to the observed values at each depth increment. For example the proportion of the error and observed value (error/observed value) for OC and AWC regardless of cluster at 0-10cm and 80-100cm were both found to be very close to 1. For AWC the distribution of model errors was largest for the extragrade cluster at all depth increments. This upholds the notion that reliability of the prediction in this part of the feature space is low in comparison to the other clusters. This relationship was mostly observed also for OC, but there were some exceptions (for example Figure 3.4.2d). Applying the penalisation calculation for the extragrade class (Tranter et al. 2010) ensured that PIC_{ej}^{L} and PIC_{ej}^{U} were larger than those of the other clusters at each depth increment.

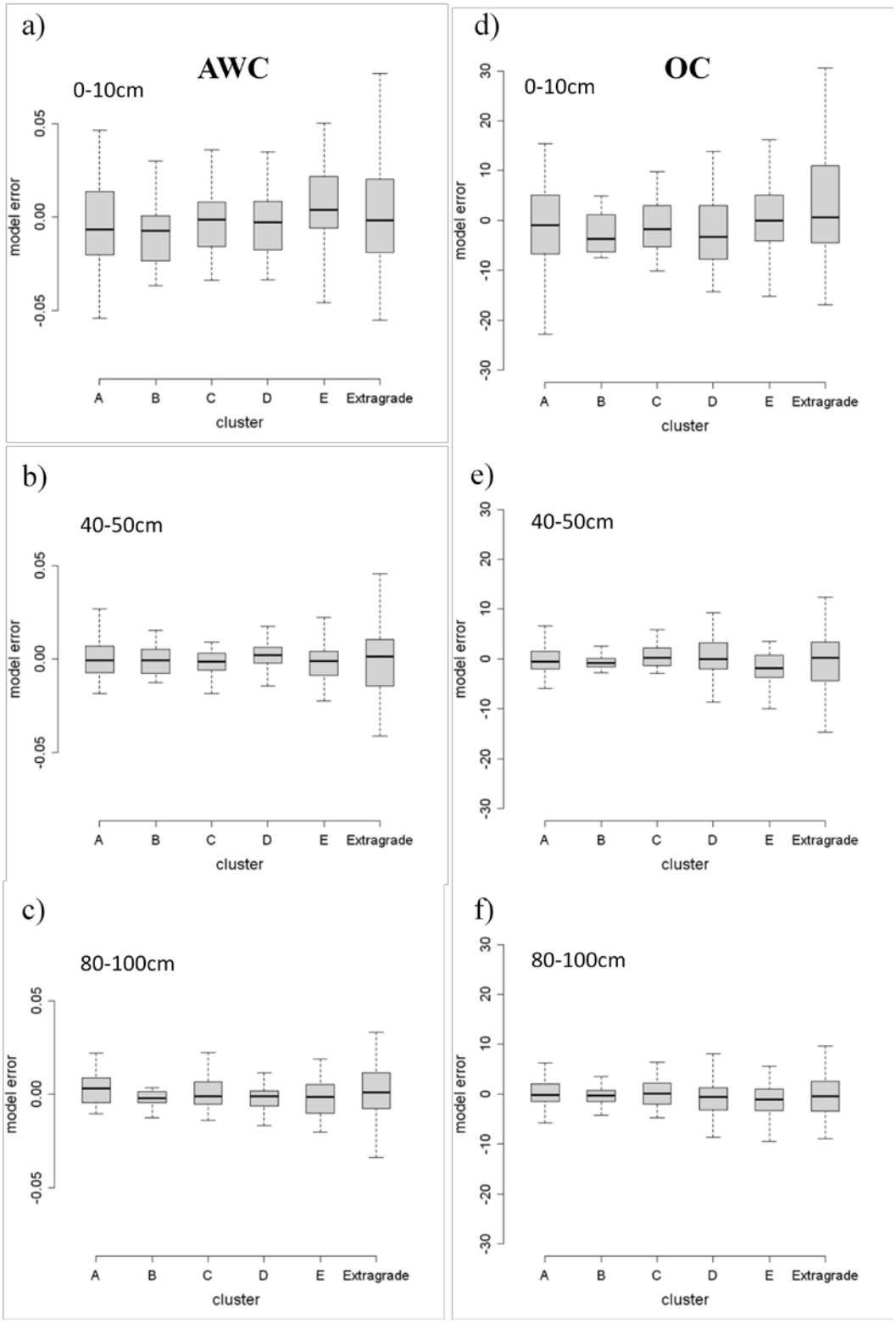


Figure 3.4.2. Box plots of the empirical distributions of model error as derived from the empirical uncertainty model at the selected depths of 0–10cm, 40–50cm and 80–100cm for AWC (a-c) and OC (d-f).

3.4.2.2. Uncertainty model validation

The final prediction and corresponding upper and lower predictions at the 8 specified depth intervals were used as parameters to construct estimated splines. This process created 3 continuous profiles for each observation: a final prediction profile and lower and upper prediction limit profiles. I randomly selected 3 validation sites to illustrate how these splines compare to the measured values of OC and AWC at each of these sites. As can be seen in Figure 3.4.3, the dotted lines constitute the final prediction, while the solid lines equate to the lower and upper prediction limits, conferring the 95% PI. The bars represent the measured values of AWC (Figure 3.4.3a-c) and OC (Figure 3.4.3d-f). As can be seen, a measured value fits within a PI if the right vertical side of the bar fits completely within the confines of the solid lines.

Due to the proportionality of the observed values with the prediction errors, the PIs are mostly widest at the soil surface. Generally for AWC, the PI then narrows gradually from about 40cm to a roughly equally spaced interval to 1m. In the example of (Figure 3.4.3b), however, the interval is narrowest at 1 m. A similar trend is evident for OC (Figure 3.4.3d-f) where uncertainty is greatest towards the soil surface. Additionally for OC there are a couple of instances where the PI does not completely confine the measured value (e.g. Figure 3.4.3f). By deriving the relative proportion of the PI (range of the upper and lower prediction limits) and the corresponding prediction of all validation observations at each depth, it was found that the proportions slightly increased with depth. This indicates an increasing uncertainty with increasing soil depth, which was also the finding when validating the prediction model. For AWC again the PI was found generally to be proportional to the prediction at a given depth; however the predictions were more accurate at the sub-soil depth increments. This was also reflected when validating the prediction model for AWC, and counters that found in Chapter 2 which could be put down to the difference in covariates and modelling steps used between the two studies.

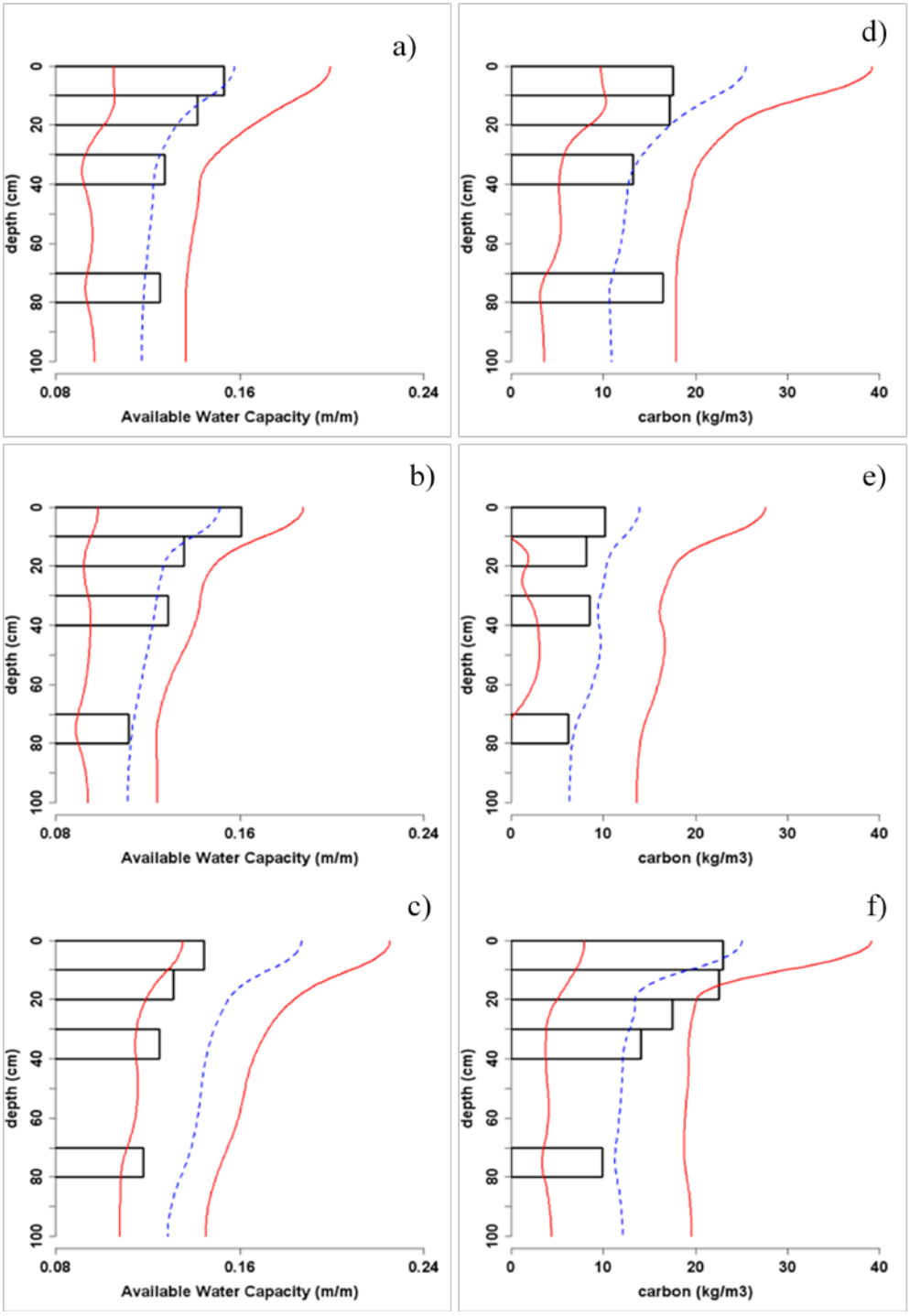


Figure 3.4.3. Profile plots of AWC (a-c) and OC (d-f) at randomly selected validation sites. Bars represent actual observed values. Dotted lines represent final DSM predictions. Solid lines represent upper and lower prediction limits.

Results of the PICP analysis indicate that at the desired confidence level of 95%, 91 and 93% of all observations fitted with their given PIs for OC and AWC respectively, indicating with this type of validation, that the empirical uncertainty model is optimal for both soil attributes (Figure 3.4.4a-b) . Furthermore, with each successive decrease in the confidence levels a near corresponding decrease in the PICP is observed for both attributes indicating a required outcome in terms of sensitivity of the PI to changing confidence levels.

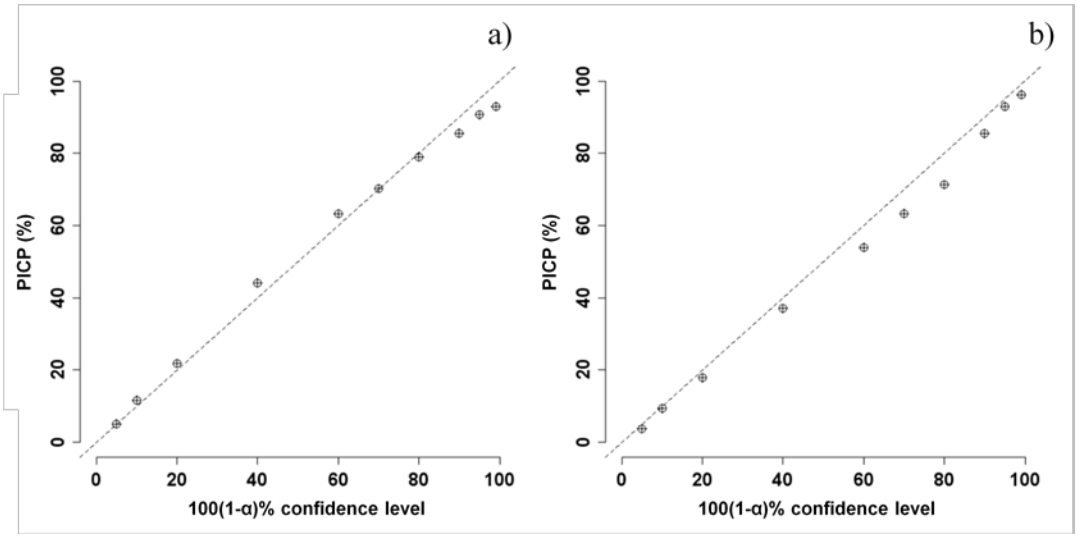


Figure 3.4.4. Prediction interval coverage probability plots (PICPs) for OC (a) and AWC (b).

3.4.3. Mapping of predictions and their uncertainties

The maps in Figure 3.4.5 illustrate the spatial variability of the degree of membership each prediction node has to each class, based on the cluster centroids derived from the empirical uncertainty model procedure. This gives a good representation of which areas in the extent of the study area share a similarity based on the given suite of environmental covariates.

Areas with a high degree of membership to the extragrade class (Figure 3.4.5a) appear to correspond to regions that are topographically diverse such as that to the east and south west where undulating slopes and hills are situated. These areas are also moderately to densely-populated by native forest. While sample sites exist across these landscapes in a predominantly equilateral triangular grid design (McGarry *et al.* 1989), it is likely that given the combination of a diverse landscape (rolling and undulating hills) and forest, these areas have not been sufficiently defined from the few sites that were taken in the area in terms of our clustering

procedure. Because I associate instances that have a high belongingness to the extragrade class as having a high prediction uncertainty, such areas could become the focus of future targeted sampling projects in order to generate new knowledge. In other instances, farm reservoirs (which if true should be eliminated from future analyses), as seen by symmetric shapes predominantly in the west of the study area, also have a high extragrade membership.

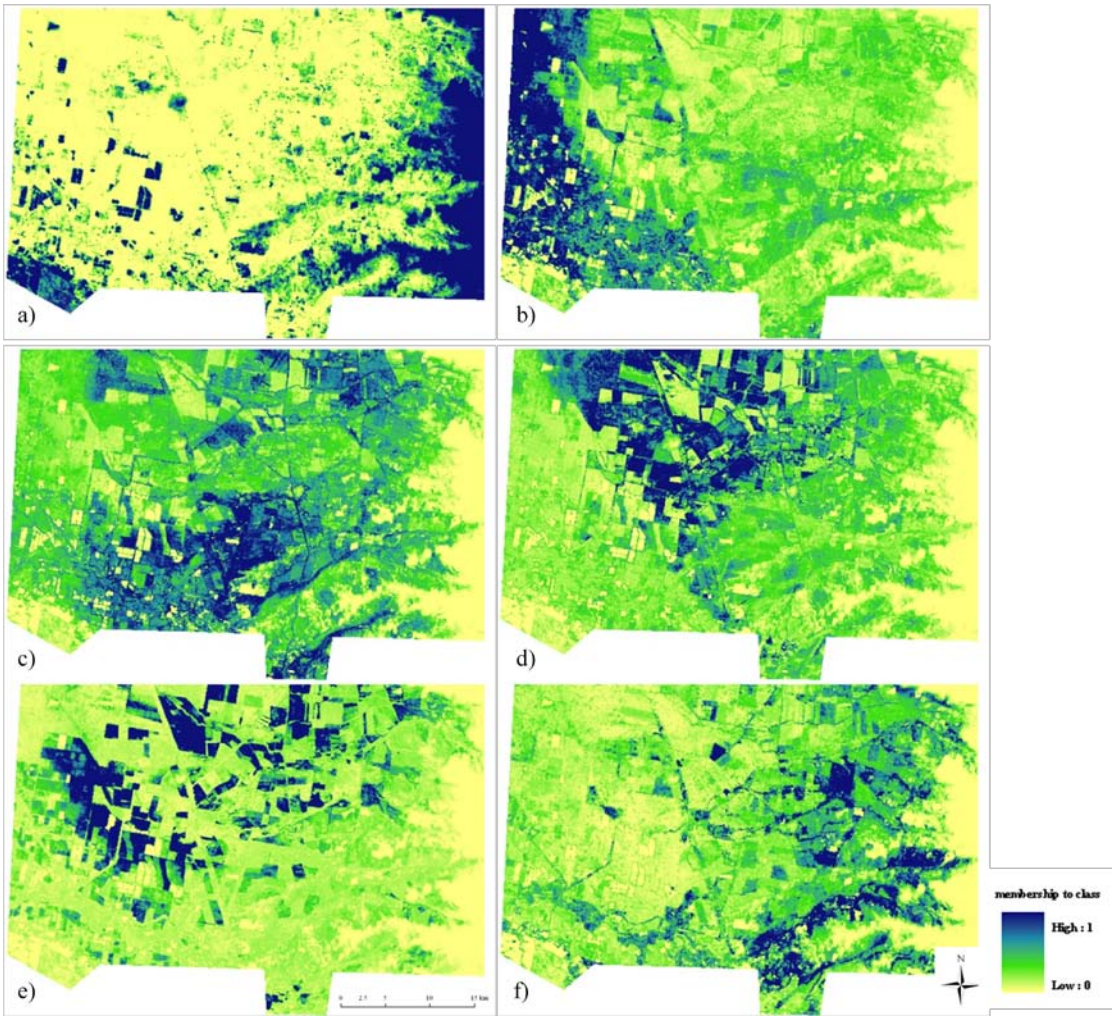


Figure 3.4.5. Spatial variation of the degree of membership each instance has to each cluster including the extragrade class. Extragrade (a), cluster A (b), cluster B (c), cluster C (d), cluster D (f), cluster E (g).

Figure 3.4.6a-c shows the variability of OC across the study area at 3 selected model depth increments of 0–10cm, 30–40cm and 80–100cm. At each depth increment there is a lower prediction limit map (Figure 3.4.6a1-c1), the final predicted map (Figure 3.4.6a2-c2) and an upper prediction limit map (Figure 3.4.6a3-c3). Similarly in Figure 3.4.7a-c**,** the spatial variability of AWC at the same depth increments as a lower (Figure 3.4.7a1-c1), final (Figure 3.4.7a2-c2), and upper prediction (Figure 3.4.7a3-c3).

Based on the lower and upper prediction limits at 0–10cm across the study area, the average concentration of OC was predicted to range between 4–35 kg m^3. The average predicted OC concentration at this depth was 18 kg m^3. At 30–40 cm the predicted average OC was 10 kg m^3. I am 95% confident that the true average of OC

at this depth is between 3–19 kg m^3. While at the 80–100cm depth increment, the predicted average of OC was 7 kg m^3, and I am 95% confident that the true average is between 2–16 kg m^3.

For AWC at 0–10cm the average was predicted to be 0.16 m/m. With 95% confidence the true average is expected to be between 0.11–0.21 m/m at this depth. At the depth increments of 30–40cm and 80–100cm the average AWC was predicted to be 0.13 (0.08–0.16) m/m and 0.11 (0.08–0.14) m/m respectively.

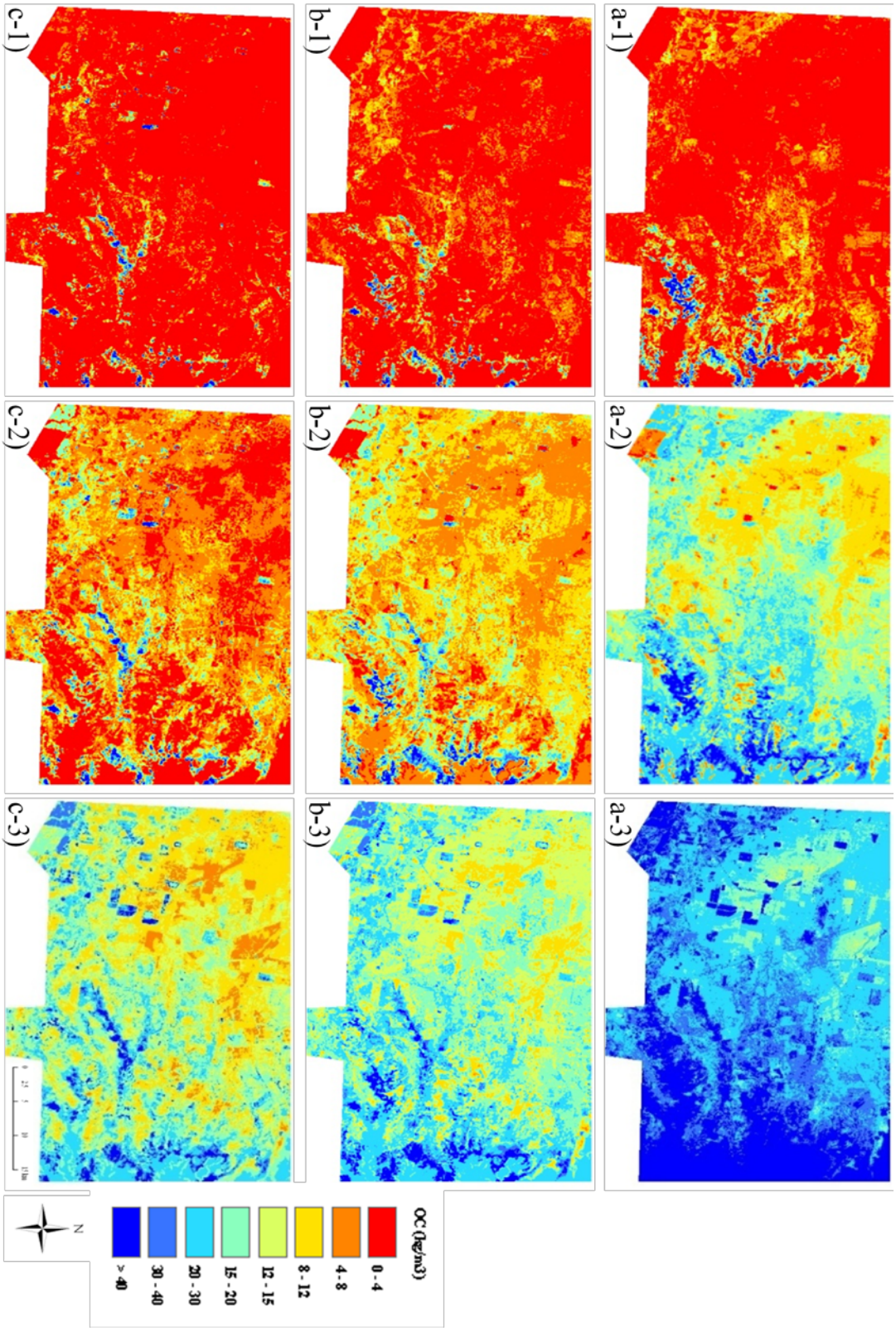


Figure 3.4.6. Variability of OC at 0–10cm, 30–40cm and 80–100cm across the Edgeroi study area. Lower prediction limit (1), DSM final prediction (2), upper prediction limit (3).

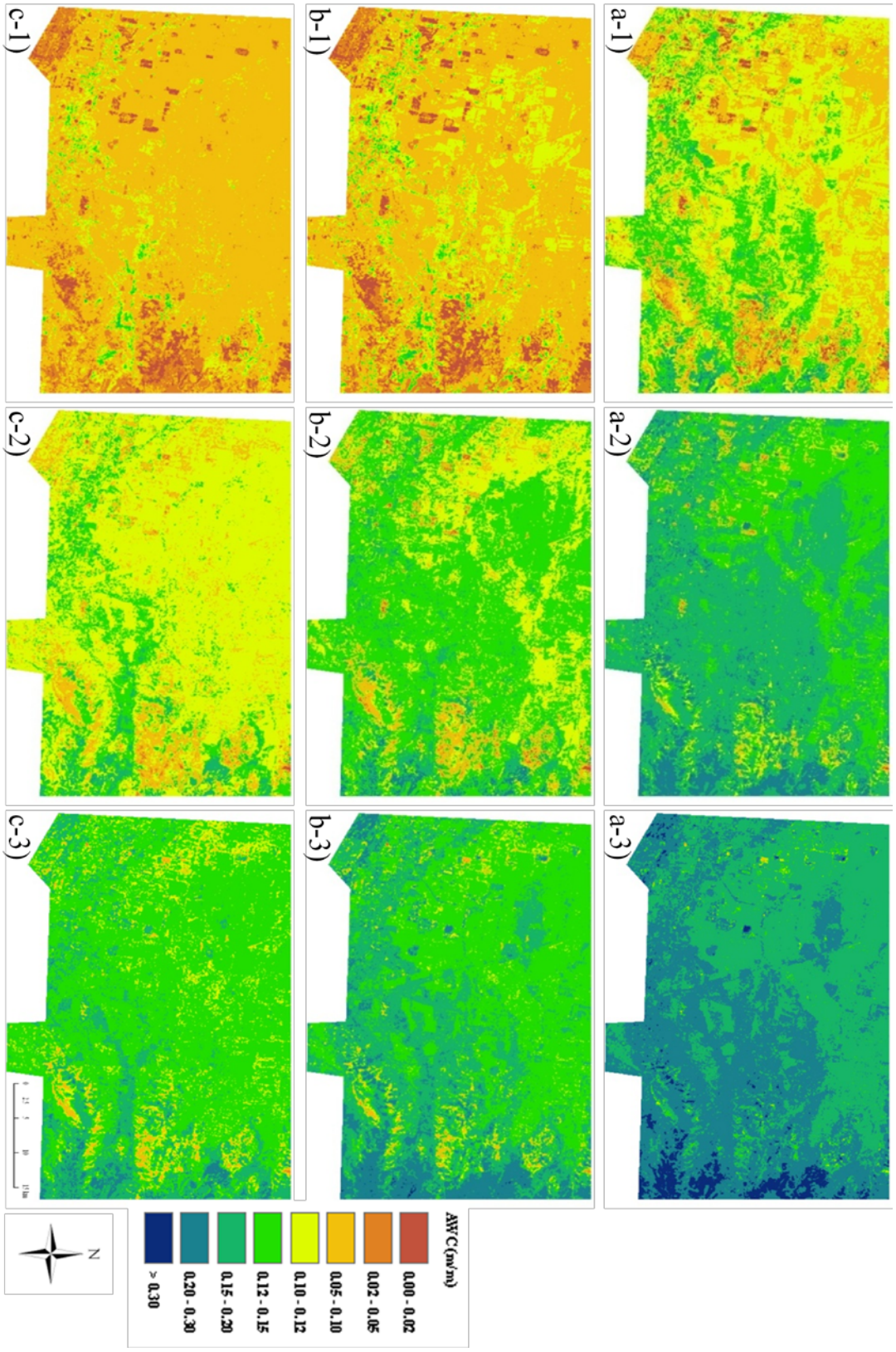


Figure 3.4.7. Variability of AWC at 0–10cm, 30–40cm and 80–100cm across the Edgeroi study area. Lower prediction limit (1), DSM final prediction (2), upper prediction limit (3).

Total AWC and OC maps to 1m are shown in Figure 3.4.8a and Figure 3.4.8b respectively. Based on these maps I predict that the average total water (m^2) to 1m is 127 (88–158) mm. The total OC estimated across the extent of the study area is predicted to be 191 (50–385) Gg in the top 1m of soil.

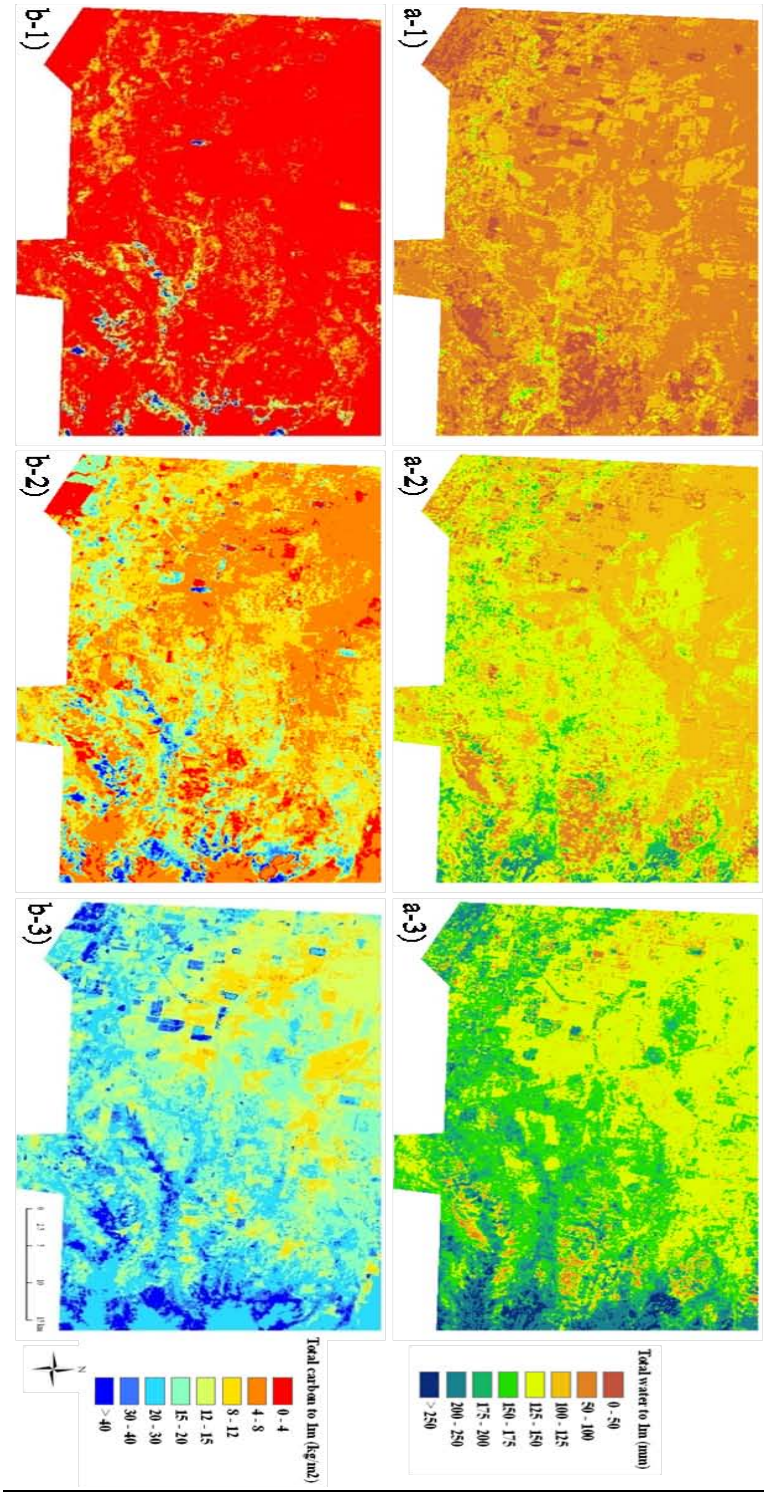


Figure 3.4.8. Total water to 1m (a) and total OC to 1m (b) across the Edgeroi study area. Lower prediction limit (1), DSM final prediction (2), upper prediction limit (3).

3.5. Conclusions

In this chapter a methodological framework was established for mapping uncertainties in the form of a PI for predicted soil attributes as they vary continuously with depth and space for a defined study area. This methodology complements the continuous prediction of soil attributes in a vertical and lateral space using splines and DSM methods. The methodology for deriving PIs is independent of the prediction model, requiring only the model outputs and the measure of error associated with those predictions

The best available quantitative measure of the deviation between the modelled output and the modelled real-world process is the residual or error. Therefore, the empirical uncertainty model explicitly accounts for all sources of uncertainty without the requirement to separate out the contribution of each error source to the overall uncertainty. I have demonstrated that this method performed well for both OC and AWC where for a given confidence level, a near matching proportion of validation observations were within the confines of the PI. While an encouraging result, I accept that this methodology represents a pragmatic approach to estimating uncertainties both spatially and laterally in a DSM framework. It is likely their estimation may be a lot more complex than that formulated in this study. Future research will obviously need to investigate the extent of this perceived complexity and the scope of future research would initially involve comparison of the empirical approach I have presented with other approaches such as Monte Carlo simulations to construct PIs. During such a comparative exercise, one would then also need to consider, in addition to the assessing the accuracy of the estimated uncertainties, the time and costs for implementing such alternative approaches.

In the course of this work however, a number of issues were presented that need to be addressed in order for improvements to be made to this approach. First, it begins with an independent model prediction framework where prior to modelling, a multivariate analysis should be performed to determine what the most closely correlated environmental covariates are for each soil attribute. This will ideally address some of the issues regarding the performance of the prediction models. Having a stronger prediction will naturally transfer to a reduced error. Ultimately, the distribution of errors for each class will be narrower, resulting in PIs that display more precision than that which I have just presented. By having an independent

modelling process also means the requirement of an independent LOCV and clustering process. This inevitably generates more work. However, while better prediction outcomes are expected, the empirical uncertainty method is neither computationally demanding nor difficult to implement.

Beyond such pertinent modifications, future work could determine how well the methodology handles both different calibration sample sizes and other soil property data. Chapter 4 investigates an independent sampling design for determining soil map quality which validates the mapped predictions *and* the quantifications of the uncertainties. On the idea of sampling, it would be ideal to investigate whether sampling in areas that have a high membership to the extragrade class would facilitate narrowing the uncertainty where the uncertainty is believed to be greatest. Methods for determining the optimal class size will also need to be investigated.

3.6. References

Altman, D.G. and Gardner, M.J., 1988. Calculating confidence intervals for regression and correlation. British Medical Journal, 296: 1238-1242.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.

Bezdek, J.C., Ehrlich, R. and Full, W., 1984. FCM- The fuzzy c-means clustering algorithm. Computers & Geosciences, 10: 191–203.

Bishop, T.F.A., McBratney, A.B. and Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, 91: 27-45.

Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A. and Selige, T., 2002. Soil regionalisation by means of terrain analysis and process parameterisation. In: E. Micheli, F. Nachtergaele and L. Montanarella (Editors), Soil Classification 2001. European Soil Bureau, Research Report No. 7, EUR 20398 EN, Luxembourg, pp. 213 – 222.

Bragato, G., 2004. Fuzzy continuous classification and spatial interpolation in conventional soil survey for soil mapping of the lower Piave plain. Geoderma, 118: 1-16.

Brown, J.D. and Heuvelink, G.B.M., 2005. Assessing uncertainty propagation through physically based models of soil water flow solute transport. In: M. Anderson (Editor), Encyclopaedia of Hydrological Sciences. John Wiley and Sons, Chichester.

Gallant, J.C. and Dowling, T.I., 2003. A multi-resolution index of valley bottom flatness for mapping depositional areas. Water Resources Research, 39:1347-1360.

Grimm, R. and Behrens, T., 2010. Uncertainty analysis of sample locations within digital soil mapping approaches. Geoderma, 155: 154–163.

Grinand, C., Arrouays, D., Laroche, B. and Martin, M.P., 2008. Extrapolating regional soil landscapes from an existing soil map: Sampling intensity, validation procedures, and integration of spatial context. Geoderma, 143: 180-190.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer, New York, N.Y.

Jain, A.K., Murty, M.N. and Flynn, P.J., 1999. Data clustering: A review. ACM Computing Surveys, 31: 264-323.

Kempen, B., Brus, D.J., Heuvelink, G.B.M. and Stoorvogel, J.J., 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. Geoderma, 151: 311-326.

Lagacherie, P., Cazemier, D.R., vanGaans, P.F.M. and Burrough, P.A., 1997. Fuzzy k-means clustering of fields in an elementary catchment and extrapolation to a larger area. Geoderma, 77: 197-216.

McBratney, A.B., 1992. On variation, uncertainty and informatics in environmental soil management Australian Journal of Soil Research, 30: 913–935.

McBratney, A.B. and de Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science, 43: 159–175.

McBratney, A.B. and Moore, A.W., 1985. Application of fuzzy-sets to climatic classification. Agricultural and Forest Meteorology, 35: 165–185.

McBratney, A.B. and Odeh, I.O.A., 1997. Application of fuzzy sets in soil science: Fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma, 77: 85-113.

McBratney, A.B., Mendonca-Santos, M.L. and Minasny, B., 2003. On digital soil mapping. Geoderma, 117: 3–52.

McGarry, D., Ward, W.T. and McBratney, A.B., 1989. Soil studies in the Lower Namoi Valley: methods and data. The Edgeroi Dataset. CSIRO Division of Soils, Adelaide.

McQueen, J., 1967. Some methods for classification and analysis of multivariate observations, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA, pp. 281–297.

Minasny, B., McBratney, A.B., 2002a. FuzME version 3.0, Australian Centre for Precision Agriculture, The University of Sydney, Australia

Minasny, B. and McBratney, A.B., 2002b. Uncertainty analysis for pedotransfer functions. European Journal of Soil Science, 53: 417–429.

Minty, B., Franklin, R., Milligan, P., Richardson, M., Wilford, J., 2009. The Radiometric Map of Australia. Exploration Geophysics 40: 325-333.

Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1992. Soil pattern recognition with fuzzy-c-means: application to classification and soil-landform interrelationships. Soil Science Society of America Journal, 56: 506-516.

Saunders, A.M. and Boettinger, J.L., 2007. Incorporating classification trees into a pedogenic understanding raster classification methodology, Green River Basin, Wyoming, USA. In: P. Lagacherie, A.B. McBratney and M. Voltz (Editors), Digital Soil Mapping: An introductory perspective. Developments in Soil Science. Elsevier, Amsterdam, pp. 389–399.

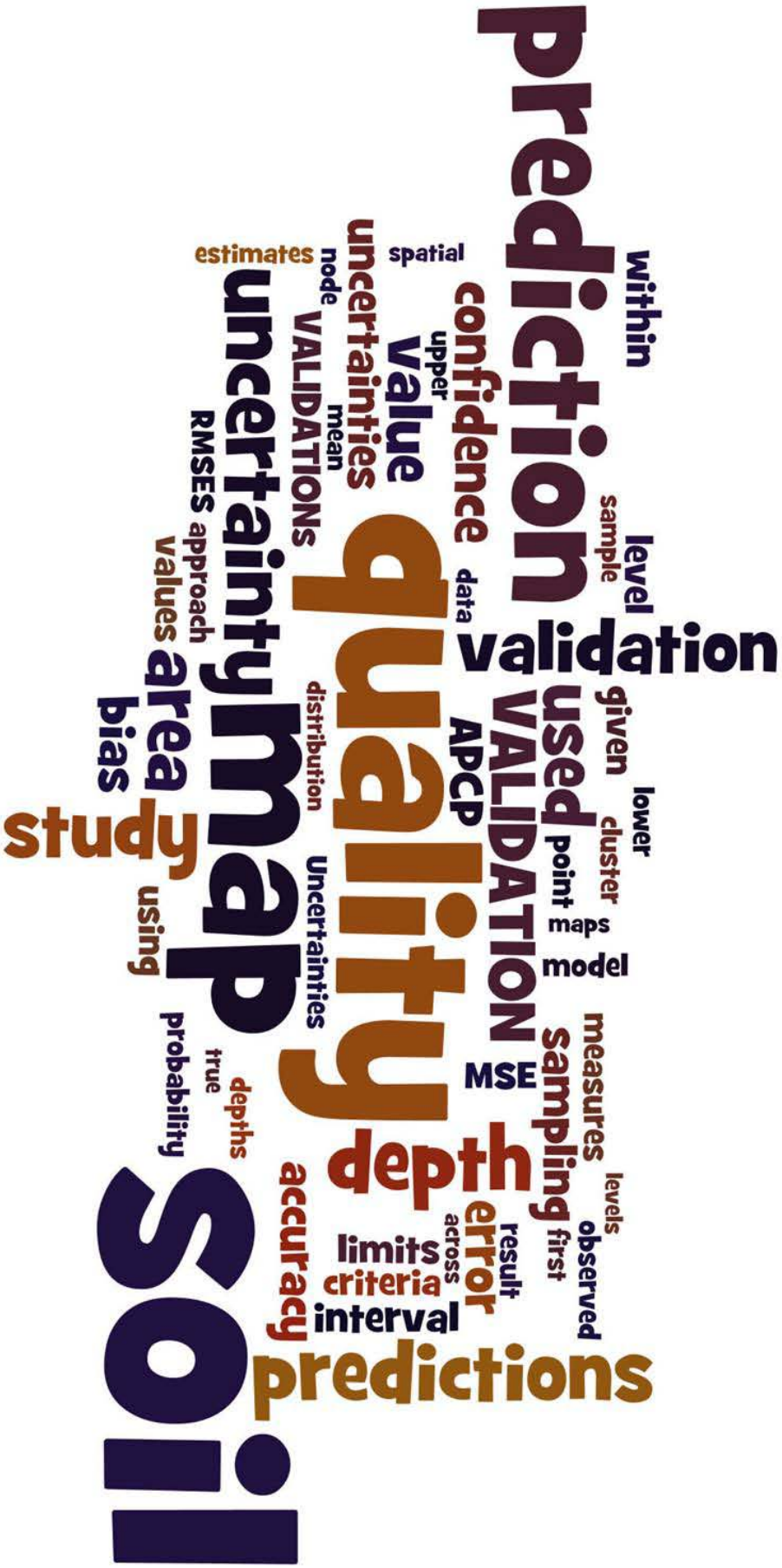
Shrestha, D.L. and Solomatine, D.P., 2006. Machine learning approaches for estimation of prediction interval for the model output. Neural Networks, 19: 225-235.

Solomatine, D.P. and Shrestha, D.L., 2009. A novel method to estimate model uncertainty using machine learning techniques. Water Resources Research, 45: Article Number: W00B11.

Tranter, G., Minasny, B. and McBratney, A.B., 2010. Estimating Pedotransfer Function Prediction Limits Using Fuzzy k-Means with Extragrades. Soil Science Society of America Journal, 74: 1967-1975.

Ward, W.T., 1999. Soils and landscapes near Narrabri and Edgeroi, New South Wales, with data analysis using fuzzy k-means. CSIRO Division of Soils Divisional Report.

Webster, R., 2000. Is soil variation random? Geoderma, 97: 149–163.


prediction
soil
quality
map
uncertainty
validation
depth
predictions
study
area
bias
value
confidence
uncertainties
VALIDATION
used
sampling
error
accuracy
limits
criteria
interval
MSE
APCP
RMSES
VALIDATIONS
estimates
spatial
within
level
sample
data
values
approach
distribution
given
lower
cluster
point
maps
model
Uncertainties
using
probability
true
depths
measures
levels
observed
first
result
across
mean
upper
node

See what the land is like and whether the people who live there are strong or weak, few or many. What kind of land do they live in? Is it good or bad? What kind of towns do they live in? Are they unwalled or fortified? How is the soil? Is it fertile or poor? Are there trees in it or not? Do your best to bring back some of the fruit of the land.
[Numbers 13:18–20 (NIV)]

Chapter 4

Criteria and sampling for simultaneously measuring the quality of predictions *and* their uncertainties in a digital soil mapping framework

Summary

In this chapter two new criteria are introduced to assess the quality of digital soil property maps. Soil map quality is estimated on the basis of validating both the accuracy of the predictions and their uncertainties (which are expressed as a prediction interval). The first criterion is an accuracy measure that is different in form to the usual Mean Square Error (MSE) because it accounts also for the prediction uncertainties. This measure is the spatial average of the statistical expectation of the Mean Square Error of a simulated random value (MSES). The second criterion addresses the quality of the uncertainties which is estimated as the total proportion of the study area where the (1-α)-prediction interval (PI) covers the true value. Ideally, this areal proportion equals the nominal value (1- α), here 95%. In the Lower Hunter Valley, NSW, Australia, I used both criteria to validate a soil pH map using additional units collected from a probability sample at five depth intervals: 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm. For the first depth interval (0-5 cm) in 96% of the area, the 95% PI of pH covered the true value. The Root Mean Squared Simulation Error (RMSES) at this depth was 1.0 pH units. Generally, the discrepancy between the expected value and the actual areal proportion in addition to the RMSES, increased with soil depth, indicating a growing imprecision of the map and underestimation of the uncertainty with increasing soil depth. In exploring this result, conventional map quality indicators emphasised a combination of bias and imprecision particularly with increasing soil depth. There is great value in coupling conventional map quality indicators with those which I propose in this chapter as they target the decision making process for improving the precision of maps and their uncertainties. For the study area I discuss options for improving upon these results in addition to determining the possibility of extending a

similar sampling approach for which multiple soil property maps can be validated concurrently.

4.1. Introduction

Quantifying soil map quality has been an area of sustained research for well over forty years, with seminal papers dating back to the 1960s and 70s e.g. Webster and Beckett (1968) and Burrough et al. (1971). Brus et al. (2011) demonstrate not only the continual development of methodologies for validating soil maps but also the importance of quantifying soil map accuracy. Grunwald (2009) however, points out that while digital soil mapping (DSM) has become popularised in recent times, it is frequently the case that maps are not validated.

Generally, soil map quality has been related to measures of accuracy (Finke 2007). These are conventionally based on measures of variance between observed and predicted values (Bishop et al. 2001). For soil property mapping, this is quantified by the goodness of fit (R^2) or the mean of the squared prediction error (MSE). For assessing the quality of categorical soil maps, measures based on mapping unit purity and user's and producer's accuracies are the most common (Congalton and Green 2009; Lark 1995). While useful, these measures are limited because it is possible only to estimate the accuracy of the predictions. While a number of tools are available that allow one to express or quantify the level of uncertainty in soil functions generated from uncertain predictions of basic soil properties (Minasny and McBratney 2002; Brown and Heuvelink 2005), rarely do we consider the quality or appropriateness of them in a practical sense. For example, given a level of uncertainty regarding a prediction of a soil attribute across a spatial extent, what implications does this have in how we interpret soil phenomena or manage the soil resource in question? Is it possible to optimise inputs such as lime for soil pH management given a prescribed level of uncertainty? It is apparent therefore that there needs to be additional measures of map quality to those conventionally reported which also take into account the quality of the estimates of prediction uncertainty for a particular soil attribute. By incorporating this additional information, we are able to consolidate our limited understanding of soil variability with reciprocal measures of soil map quality.

A useful and empirical approach to estimating the uncertainty of model outputs was proposed by Shrestha and Solomatine (2006) where uncertainty was expressed in the form of two quantiles (constituting a prediction interval) of the underlying distribution of prediction errors (residuals). The prediction interval (PI) explicitly

takes into account all sources of uncertainty and circumvents attempts to separate out the contribution of each source (Shrestha and Solomatine 2006; Solomatine and Shrestha, 2009). In Chapter 3 the empirical uncertainty approach of Shrestha and Solomatine (2006) was adapted within a DSM framework to map PIs across a study area. To that end, the soil map quality criteria I propose in this study are largely based on the empirical coverage of PIs.

In terms of quality measures for validating soil maps, Brus et al. (2011) provide a comprehensive review of those used for both categorical and quantitative digital soil maps. Map validation by probability sampling involves the random selection of additional test units (observations at locations not used for model calibration) from a study area. In probability sampling all units within a study area have a positive probability of being selected; where the probabilities are determined by the sampling design and can be derived from this design (de Gruijter et al. 2006). Brus et al. (2011) concluded that probability sampling is the more superior validation method (in comparison to random holdback or cross-validation) because unbiased estimates of the measures of soil map quality can be obtained by 'design-based' inference and thus are free of model assumptions (Brus and de Gruijter 1997). This is generally not the case for random holdback or cross-validation (Brus et al. 2011). However from the point of practicability it is necessary to weigh up whether a more statistically valid estimate of map quality is worth the time, cost and effort to perform independent probability sampling.

Chapter 3 demonstrates a procedure for validation of digital soil maps (predictions and uncertainties) using a randomly held-back sample. In this chapter validation is performed on the basis of a probability sample to collect additional sampling units from the study area. Subsequently, the aim of this research therefore is to present and illustrate two new criteria (in addition to those conventionally reported) for evaluating both the quality of predictions and the quality of quantifications of the uncertainty in a DSM on the basis of 'design-based' inference.

4.2. Soil map to be validated

4.2.1. Study Area

The area selected for this study is an approximately 220 km^2 area north of the town of Cessnock (32.83°S 151.35°E) in the Lower Hunter Valley, approximately 140 km

north of Sydney, NSW, Australia (Figure 4.2.1). Topographically, this area consists mostly of undulating hills that ascend to low mountains to the south-west. This area is part of the Sydney Basin where parent materials are composed mostly of Mesozoic sandstones and shales (Thackway and Cresswell 1995). The dominant soil types according to the Australian Soil Classification (Isbell 2002) are Red and Brown Dermosols and, on topographic rises, Red Calcarosols (Odgers et al. 2011). In terms of landuse, dryland agricultural grazing systems are predominant, followed by an expansive viticultural industry. While most of the land has been dedicated for these uses, tracts of remnant natural vegetation (dry forest) are apparent, particularly towards the south-western (Broken Back Range), eastern (Werakata National Park) and northern margins of the study area. See Bell (2004) for further details of the environmental setting of this area.

For this study I chose soil pH as the target variable. The reason for this selection was because viticulture is a widespread industry across the study area and many management decisions are centred on the nutrient status of the high value wine grape crops, which is generally determined by soil pH (White 2003). Soil nutritional status affects all parts of the grape vine, from root growth and distribution through to shoot growth and grape composition.

The soil dataset I used contains 994 sites where pH was recorded at each horizon and/or at specific depths to at least 1m. Three hundred of the soil samples were collected in 2004 using the Conditioned Latin Hypercube sampling method (Minasny and McBratney 2006) where compound topographic index, parent material and Normalised Difference Vegetation Index were used as auxiliary variables to provide environmental information. The remaining 694 samples were collected sporadically between 2001 and 2009 mainly using a *k*th-order random toposequence sampling design (Odgers et al. 2008).

Figure 4.2.1. Lower Hunter Valley study area with respect to location in New South Wales (large box) and Australia (small box).

4.2.2. Digital soil mapping of the predictions

A soil map depicting the lateral and vertical distribution of soil pH across the study area was generated following the procedure of Chapter 2 which uses an amalgam of soil depth spline functions and DSM techniques (Figure 4.2.2b) (only 0-10 cm shown). The spatial entity of the map is point support where point estimates were made on a 25m regular grid. The vertical resolution at each grid node is 1cm.

This map was generated using a regression kriging approach where the predictions were based on the calibration dataset of 994 soil profile descriptions distributed across the area (Odgers et al. 2011). To standardise prediction depths, mass preserving splines were fitted individually to each profile before mean estimates of pH were taken at 0-10 cm, 10-20 cm, 20-30 cm, 30-40 cm, 40-50 cm,

50-70 cm, 70-80 cm, 80-100 cm. The deterministic component (of regression kriging) used a neural network where the target variables were the mean observations at each of the standardised depths. These were modelled against a suite of environmental covariates derived from 25m rasters of a digital elevation model (DEM) and various derived terrain attributes; Landsat 7 ETM+ band data and various band derivatives. The model formulae derived from this procedure were then applied across the extent of the study area, where covariate information existed only.

For the stochastic component, model residuals were independently mapped for each standardised depth with kriging using localised variograms (exponential function) of the 100 nearest neighbours to a prediction point. Both the deterministic and stochastic components were summed to arrive at a final prediction for each standard depth at each point.

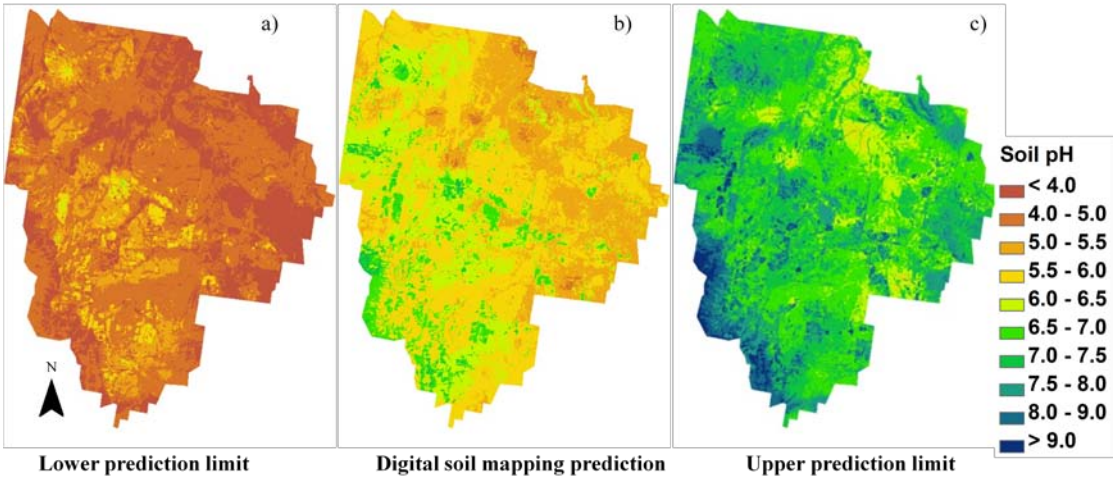


Figure 4.2.2. Soil pH map to be validated (displaying only the 0-10 cm interval). Soil map depicts the (b) digital soil mapping prediction and the (a) upper and (c) lower prediction limits which constitute a prediction interval.

4.2.3. Digital soil mapping of the uncertainties

The method described for quantifying the uncertainties for DSM in Chapter 3 was also applied in this study. This method is an adaption of the ‘‘uncertainty estimation based on local errors and clustering’’ (UNEEC) method, described by (Solomatine and Shrestha 2009). The purpose of the UNEEC is to derive the upper and lower prediction limits based on the model error, and since it is estimated through an empirical distribution, it is not necessary to make any assumption about residuals (Solomatine and Shrestha 2006).

Using the calibration data, the adaption of the UNEEC for DSM first uses fuzzy k-means with extragrades to partition the environmental covariate data into clusters which share similar environmental attributes (McBratney and de Gruijter 1992). Each observation in the dataset is then given a grade of membership to all clusters including the extragrade cluster.

Next, model output error is determined. The regression kriging approach similar to that used for the DSM prediction procedure was used except final predictions were evaluated by leave-one-out-cross-validation (Hastie et al. 2009). The only output required here is the residual between the observed value of the target variable and the corresponding final prediction at each calibration point.

Once the clusters and model output error have been identified, the PIs for each cluster are computed from the empirical distributions of the corresponding residuals.

Therefore, one underlying assumption of the empirical uncertainty method is that particular areas (clusters) within a landscape will have similar residuals or distribution of residuals and ultimately share a similar range of uncertainty. To construct a cluster 95% PI, first I assign each calibration point to the cluster it has the highest membership to. Then for each cluster, the 2.5 and 97.5 percentile values are taken from the empirical distribution of residuals for the lower and upper prediction limits respectively. Those points that belong to the extragrade cluster are handled differently in that I impose a penalisation by way of a multiplier, as proposed by Tranter et al. (2010):

$$PIC_{ej}^{L} = 2 \times q_{2.5}$$
$$PIC_{ej}^{U} = 2 \times q_{97.5}$$

[4.2.1]

where PIC_{ej}^{L} and PIC_{ej}^{U} are the lower and upper prediction limits of the extragrade class, and q is the quantile value of the extragrade cluster error distribution at each depth increment.

Computing the PI for each prediction node across the study area first requires each node to be assigned membership grades to each of the clusters characterised in the calibration procedure. These are determined on the basis of the environmental covariates at each node and the pre-determined cluster centroids. From this, I determine the prediction node PI using the weighted mean of the PI of each cluster (Shrestha and Solomatine 2006). This can be defined mathematically as:

$$PI_{j}^{L} = \sum_{i=1}^{c} m_{ij}\ PIC_{j}^{L}$$
$$PI_{j}^{U} = \sum_{i=1}^{c} m_{ij}\ PIC_{j}^{\mathrm{U}}$$

[4.2.2]

where PI_{i}^{L} and PI_{i}^{U} correspond to the weighted lower and upper PI for the *i*th prediction node, PIC_{j}^{L} and PIC_{j}^{U} are the lower and upper PIs for each cluster *j* (including extragrade cluster) as determined from the calibration sites, and m_{ij} is the membership grade of *i*th prediction node to cluster *j*. Finally, the lower and upper prediction limits (PL_{i}^{L} and PL_{i}^{U} respectively) are then derived for each prediction node by adding the prediction (from the DSM procedure) to PI_{i}^{L} and PI_{i}^{U}.

All the steps after the clustering process are repeated for each of the standardised depths. Figure 4.2.2a and Figure 4.2.2b show the lower and upper prediction limits of pH across the study area for the 0-10cm depth increment. Mass preserving splines are also used to construct continuous representations of the upper and lower prediction limits of pH by using the estimated PL_i^U and PL_i^L values at each depth respectively as parameters.

4.3. Material and Methods

4.3.1. Concepts of Soil Map Quality

In this study I use a design-based approach to derive external accuracy estimates of soil map quality in terms of both the predictions and their uncertainties. This involves the use of additional data collected from a probability sample for which the specifics will be discussed later in more detail. Using such a design-based approach means that our estimates of map quality are model-free and un-biased (de Gruijter et al. 2006; Brus et al. 2011).

For this study, the mapped predictions and their uncertainties are on point support. This means that the additional independent sampling units should also have point-support. As the first criteria of our proposed soil map quality indicators I use a measure which assesses the accuracy of predictions (but also taking into account the uncertainties). As discussed previously, the usual measure for the accuracy of quantitative map predictions is the mean squared prediction error (MSE) which is defined as:

$$MSE = \frac{1}{\|A\|}\int_{s\in A}\{z_p(s) - z(s)\}^2 \mathrm{d}s \qquad \textbf{[4.3.1]}$$

where A is the mapped area and $z_p(s)$ and $z(s)$ are the predicted and true values at location s of the target variable, here pH. Additionally bias, expressed as the Mean Error (ME) can then also be derived, which has the general formula:

$$ME = \frac{1}{\|A\|}\int_{s\in A}\{z_p(s) - z(s)\}\mathrm{d}s \qquad \textbf{[4.3.2]}$$

These conventional measures quantify only the error and bias of the predictions, it does not account for the uncertainty of the local predictions. To this end I introduce the 'Mean Squared Error of Simulation' (MSES). The underlying idea can be simply explained by imagining that a value is simulated randomly at a given location, using the local prediction and the local prediction variance given by the map as parameters of a local error distribution. The measure that I propose is the spatial average of the statistical expectation of the mean squared error of a simulated random value:

$$MSES = \frac{1}{\|A\|} \int_{s \in A} E_s\{z_s(s) - z(s)\}^2 \mathrm{d}s$$

[4.3.3]

where $z_s(s)$ and $z(s)$ are the simulated and true values, respectively, at location s and E_s is the statistical expectation over the error distribution. For estimation of the MSES, it is equated as the summation of two components; a spatially averaged (squared) local bias component, identical to the MSE and a spatially averaged local precision component:

$$\begin{aligned} MSES &= \frac{1}{\|A\|} \int_{s \in A} E_s\{z_p(s) - z(s)\}^2 \mathrm{d}s + \frac{1}{\|A\|} \int_{s \in A} E_s\{z_p(s) - z_s(s)\}^2 \mathrm{d}s \\ &= MSE + \frac{1}{\|A\|} \int_{s \in A} \sigma^2(s) \mathrm{d}s \\ &= MSE + \overline{\sigma^2} \end{aligned}$$

[4.3.4]

where $\sigma^2(s)$ is the variance of the prediction error as given by the map at location s. Note that the second component can be simply obtained by averaging the of σ^2 over the entire map grid. An estimate of MSES can therefore be obtained by estimating MSE from the sample data, and adding $\overline{\sigma^2}$ to it:

$$\widehat{MSES} = \widehat{MSE} + \overline{\sigma^2}$$

[4.3.5]

The second criterion is expressed as the areal proportion of the mapped area where the measured value at a specified depth interval fits within the bounds of its

estimated (1-α) % PI (the PI covers the measured value), shortly referred to hereafter as the areal proportion correctly predicted (APCP):

$$APCP = \frac{1}{\|A\|} \int_{s \in A} i(s) \mathrm{d}s \qquad \textbf{[4.3.6]}$$

where $i(s)$ equals 1 if the true pH value at location s is covered by the PI given by the map at s, and 0 otherwise.

Conceptually, both criteria form an agreeable pair and ideally should be reported together. The first explicitly deals with the accuracy of the simulations, while the second signals either an under- or over- coverage (PIs too narrow or too wide) for a given confidence level. As a case in point and the necessity for reporting on both criteria; it is possible to have an inaccurate map (wide prediction intervals) with the same coverage proportion (APCP) as an accurate map. Alone, the second criterion can not tell the difference between the two. Therefore coupling the APCP with the MSES allows penalisation to be given where predictions are found to be inaccurate.

4.3.2. Probability Sampling Design

A stratified simple random sampling design (STSI) (de Gruijter et al. 2006) was used to select the sampling locations at which soil pH was to be laboratory analysed at specified depth increments. This design was chosen over a simple random sample (SI) design on the basis that greater efficiency can be expected in terms of smaller sampling variance of the estimated map quality measures from the same number of samples (Brus et al., 2011). I used two stratification variables: the depth-averaged whole-profile prediction of pH and an uncertainty measure; the depth-averaged whole-profile difference between the upper and lower prediction limits. The averaged prediction and the uncertainty measure for each mapped point location were plotted on a graph from which four equal-area strata were generated empirically by shifting the threshold values of each stratification variable (Figure 4.3.1a). The total number of prediction nodes equalled 353 316 making the size of each stratum 88 329 nodes. The characteristics or threshold values for each stratum are summarised in Table 4.3.1. Figure 4.3.1b shows the spatial extent of the strata across the study area where it is worth noting that the strata do not form contiguous areas.

From knowledge of the soil landscape across the study area, the strata which have the highest uncertainty (C and D; widest PIs) appear to cover areas of mountainous terrain or rugged terrain relative to other areas. A low density of prediction sites coupled with complex terrain and for the most part inaccessible, it is quite logical that uncertainty is high in these landscape settings. Conversely, the strata that have a lower uncertainty (A and B) exist predominantly where the relief is less complex, the land has been cleared and where most agricultural pursuits are concentrated in the areas such as grazing and viticultural production. For example, stratum B, is a good indicator of where viticulture is practiced as the soils are a little higher on the landscape and are generally characterised by higher pH levels.

Table 4.3.1 Threshold values determined empirically of the stratification variables:- depth-averaged whole-profile prediction of pH and depth-averaged whole-profile difference between the upper and lower prediction limits for each stratum.

Stratum	pH prediction	Uncertainty (95% PI)
A	Low (≤ 5.9)	Narrow (≤ 2.6)
B	High (>5.9)	Narrow (≤ 2.6)
C	Low (≤ 5.9)	Wide (>2.6)
D	High (>5.9)	Wide (>2.6)

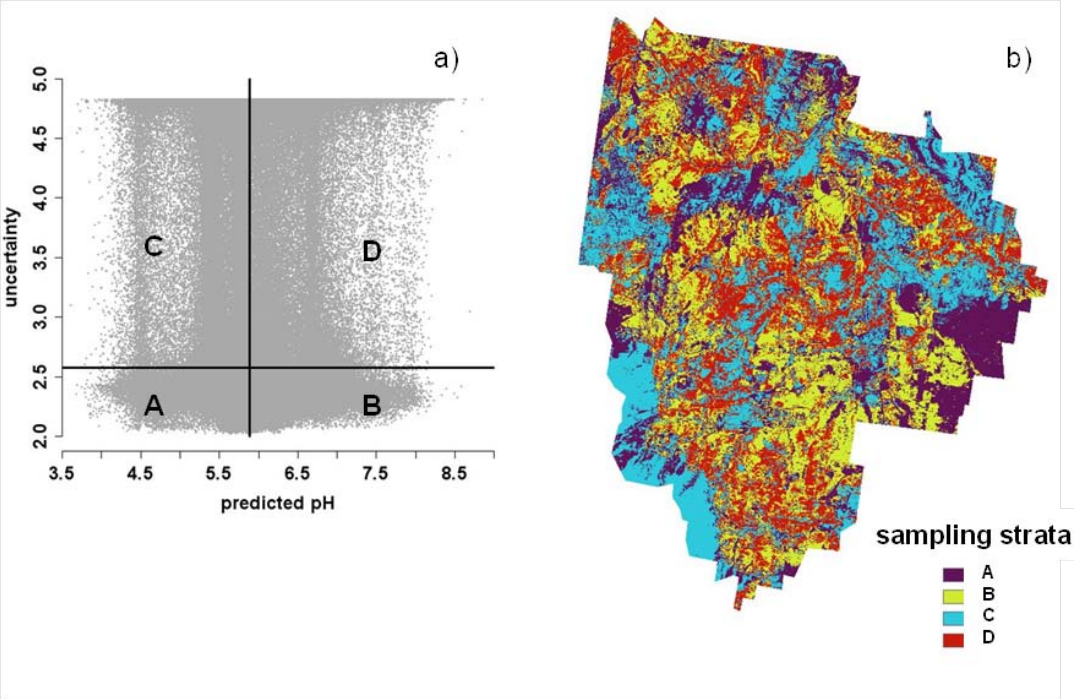


Figure 4.3.1. Determination of sampling strata and their subsequent spatial extent. (a) Plot illustrating the process for constructing equal-area strata where the stratification variables were the depth-averaged whole-profile pH prediction and depth-averaged whole-profile difference between upper and lower prediction limits. Black lines indicate the threshold values for demarcation of each stratum A, B, C and D. (b) Spatial coverage of the equal-area strata across the study area.

A total of 100 sampling units (SUs) were used to validate the map in this study. These were allocated proportionally to the strata, so that from each stratum 25 nodes were selected fully randomly from the potential 88329 nodes. A handheld Global Positioning Satellite (GPS) receiver was used to locate the positions of the SUs within the field. The SUs were soil cores of between 100-120 cm in length and a diameter of 5 cm. These were taken using a hydraulic geoprobe soil corer mounted on the back of a truck/all terrain vehicle (ATV).

In the laboratory, each SU was sub-sampled corresponding to the depth intervals of: 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm. Once mixed, a small aliquot from each depth interval was analysed using the 1:5 soil:water suspension method to determine soil pH (Rayment and Higginson, 1992). Randomly selected, duplicate aliquots totalling 50 were also analysed in order to estimate measurement error for reasons described further on.

4.3.3. Design-based Estimation of Soil Map Quality

The statistical inference for estimating the areal proportions and the standard errors of point locations which fit within their PI is given in de Gruijter et al. (2006). The indicator value for each SU at each depth was evaluated by:

- First, the PL_i^L and PL_i^U estimates corresponding to each of the 8 prediction depths were used as the spline parameters for the construction of continuous representations of the prediction limits;
- The continuous PL_i^L and PL_i^U estimates were then queried to derive the mean of the prediction limits for each of the 5 sampling depths at the validation points;
- Analysis was performed to determine whether the observed pH value fitted within its corresponding PI. Indicator values of either a one (1) indicating a fit within the bounds of the PI, or zero (0), indicating a non-fit not fit within the bounds of the PI.

From Equation 4.3.6 the design-based estimator for APCP in this study is:

$$\widehat{APCP} = \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h} I_{hj} \qquad \textbf{[4.3.7]}$$

where $\|A_h\|$ and n_h are the surface area and sample size (number of selected validation points, here 25) of stratum *h*, respectively, and I_{hj} is the indicator value as determined at the *j*-th sample point of stratum *h*. The reason why I use I_{hj} instead of i_{hj} is because I do not possess *true* values of pH and thus need to work from the *measured* pH values, meaning that the indicator values are thus subject to random error. The measurements that are used for validation may be accurate enough that one can safely assume that the effects of measurement error on the validation results are negligible. However, that assumption cannot be made in this study; therefore we need to consider the effects of measurement error. Contained within Appendix 4.1 is an explanation of accounting for measurement error in map validation. Also in Appendix 4.1 are the detailed calculations for estimation of the error variance of $\widehat{APCP}$ due to both the sampling *and* measurement errors.

For estimation of the MSES, Equation 4.3.5 is followed. As only $\widehat{MSE}$ is subject to sampling error, the sampling variance and standard error of $\widehat{MSES}$ equal those of $\widehat{MSE}$. Finally, an estimate at the scale of the mapped variable, the root mean squared

error of simulation (RMSES), is obtained by taking the square root of $\widehat{MSES}$. Similarly for the limits of a confidence interval.

The methodology for estimation of the APCP is evaluated on the basis of a 95% PI. One way to assess both the performance and sensitivity of the PIs is to estimate the APCP using a range of confidence levels (besides 95%). Variants of this type of validation, termed the prediction interval coverage probability (PICP) have previously been shown to be an important validation criterion in other studies e.g. Shrestha and Solomatine (2006) and Tranter et al. (2010). Akin to the estimation of a 95% PI, I performed the same procedures previously described to construct PIs with confidence levels ranging from 5%–90% and 99%. For each confidence level, estimation of the APCP was derived in the same way as described for a 95% confidence level. The PICP is a valuable indicator of the validity of the uncertainty model where it is said to be optimal when the PICP value is close to the range of corresponding $100(1-\alpha)\%$ confidence levels.

4.3.4. Conventional Measures of Map Quality

For conventional measures of map quality, I compared the predicted values of pH at each depth for each SU with the corresponding observed value. The method (using mass-preserving splines) for evaluating the prediction at each sampling depth was the same as for deriving the lower and upper prediction limits.

I derived spatial estimates of accuracy, bias and imprecision at each depth interval, where accuracy is stated in terms of the Root Mean Squared Error (RMSE) which is estimated by taking the square root of the predicted MSE (Equation 4.3.1). As I cannot assume that the measurements used for validation are free of the effects of measurement errors, it is necessary to make allowances for them within the calculation. Appendix 4.1 provides unbiased estimates for the MSE and variance of MSE in the presence of measurement errors. Confidence limits for the RMSE can be calculated as the square roots of the confidence limits for MSE, whereby $\widehat{MSE}$ is assumed normally distributed. For calculating the variance of the ME (bias) there is no need to correct for measurement error, because no non-linear transformation is involved which means it is automatically accounted for in the standard estimation of the variance.

Imprecision was simply calculated as the square root of the difference between the $\widehat{MSE}$ and the squared $\widehat{ME}$ defined as:

$$IMP = \sqrt{\widehat{MSE} - \left(\widehat{ME}\right)^2}$$

[4.3.8]

4.3.5. Implementation of methods

See section 3.3.5 for details of softwares used in this study.

4.4. Results and Discussion

As a preamble to interpreting the output of the proposed map quality criteria, there are four general outcomes to consider. The first is that the APCP is equal to or very near (above or below) the nominal confidence level and the RMSES is small. For this result the predictions and simulations are accurate but more importantly, what uncertainty there is, is adequately accounted for and covered by the PI. This result is ideal, however the map user will need to consider the level of uncertainty they are willing to accept for the given purpose of the map. In the case of this study, the purpose may be the optimisation of liming where viticulture is practiced. Nevertheless, given that the accuracy is quite high for this first outcome, logic will generally indicate that the uncertainties will also be quite low (narrow prediction intervals). The second outcome is the situation where the APCP is as the first but the RMSES is large. In this result the mapped predictions are inaccurate, but the PIs manage to account for most if not all the known sources of uncertainty. This is not an ideal result, because ultimately the map user wants an accurate map and the PIs could be considered to be too wide, thus making precise management decisions difficult, for example, the optimisation of lime inputs. It should be avoided, however, to interpret a map as being either high or low quality without consideration of the intended purpose of the map. Nevertheless, this second result is more ideal than the third whereby the difference between the APCP and the nominal confidence level is large, and the RMSES is high. This result indicates an inaccurate map in addition to the presence of other sources of uncertainty that were not accounted for and/or bias. The fourth outcome is where the RMSES is small but the APCP is far from the nominal confidence level. Given the occurrence of any one of the four outcomes, it then becomes necessary to investigate reasons why for example the RMSES is large but the APCP is ideal (second outcome) etc. It is therefore worth stressing at this

point that conventional map quality indicators should also be used as a companion to our proposed criteria because they address explicitly issues of bias and imprecision.

In terms of the results from this study, the PI was computed for a 95% confidence level. With this basis, the ideal APCP for any of the specified depth intervals should be 95%. Taking into account the standard errors of the estimated areal proportions, this condition was met at the 0–5 cm and 5-15 cm depth intervals where proportions of 96 ± 4% and 88 ± 7% were estimated, respectively (Table 4.4.1). This means that at least to 15 cm, I am 95% confident the true value of pH falls within the specified PI at each prediction node. The RMSES at these depths was found to be 1.0± 0.1 units of pH, indicating that there may be some issues of accuracy. However, relative to the RMSES values from 30 cm, the predictions towards the soil surface are more accurate. This issue of inaccuracy is corroborated with the conventional map quality indicators where it can be seen just at the first two depth increments that there is evidently some positive bias (indicating systematic under prediction) within the predictions. More detail of the conventional indicators is discussed further on.

Table 4.4.1. Results of the proposed soil map quality indicators: The areal proportion of the map within the specified prediction interval or correctly predicted (APCP) and Root mean square error of simulation (RMSES) at each depth increment. Additionally, corresponding measures of accuracy (RMSE), bias (ME) and imprecision (IMP) of the given map taking into account only the quality of the predictions.

Depth (cm)	APCP	RMSES	RMSE	ME	IMP
0-5	96 ± 5%	1.0± 0.1	0.6± 0.1	0.2± 0.1	0.6
5-15	88 ± 7%	1.0± 0.1	0.7± 0.1	0.3± 0.1	0.6
15-30	81 ± 8%	1.0± 0.1	0.8± 0.1	0.2± 0.1	0.7
30-60	74 ± 9%	1.3 ± 0.1	1.0± 0.1	0.1± 0.1	1.0
60-100	81 ± 8%	1.6 ± 0.1	1.1± 0.1	0.0± 0.1	1.1

With increasing depth down the profile an underestimation of the uncertainty is observed. For example, at 15-30 cm the APCP indicates that on average 81± 8% of map nodes have PIs that cover the true value of pH at that depth interval. At 30-60 cm the APCP decreased to 74± 9%, then increased marginally at the 60-100 cm depth interval (81± 8%). At these three depth increments the RMSES increased from 1± 0.1, to 1.3± 0.1 and finally 1.6± 0.1, respectively. Taken as a whole, the picture

that is quite evident is that with increasing soil depth, the accuracy of the predictions decreases along with a growing level of uncertainty that is not accounted for.

Observed vs. fitted plots provide a visual guide of the deviation of the predicted values of pH from the co-located measured values at each depth interval (Figure 4.4.1). Lin's concordance correlation coefficient (CCC) was used to derive a quantitative measure of the relationship between the predicted and observed measurements (as described in Chapter 2- 2.2.5.) The plots (Figure 4.4.1) show a moderate agreement between the observed and fitted values where CCC ranged between 0.44 and 0.30, with the strongest predictions at 0–5 cm (Figure 4.4.1a). Similarly the accuracy at 0–5 cm was 0.6± 0.1 and gradually decreased with depth to 0.7± 0.1 (5–15 cm), 0.8± 0.1 (15–30 cm), 1.0± 0.1 (30–60 cm) and 1.1± 0.1 (60–100 cm)**.**

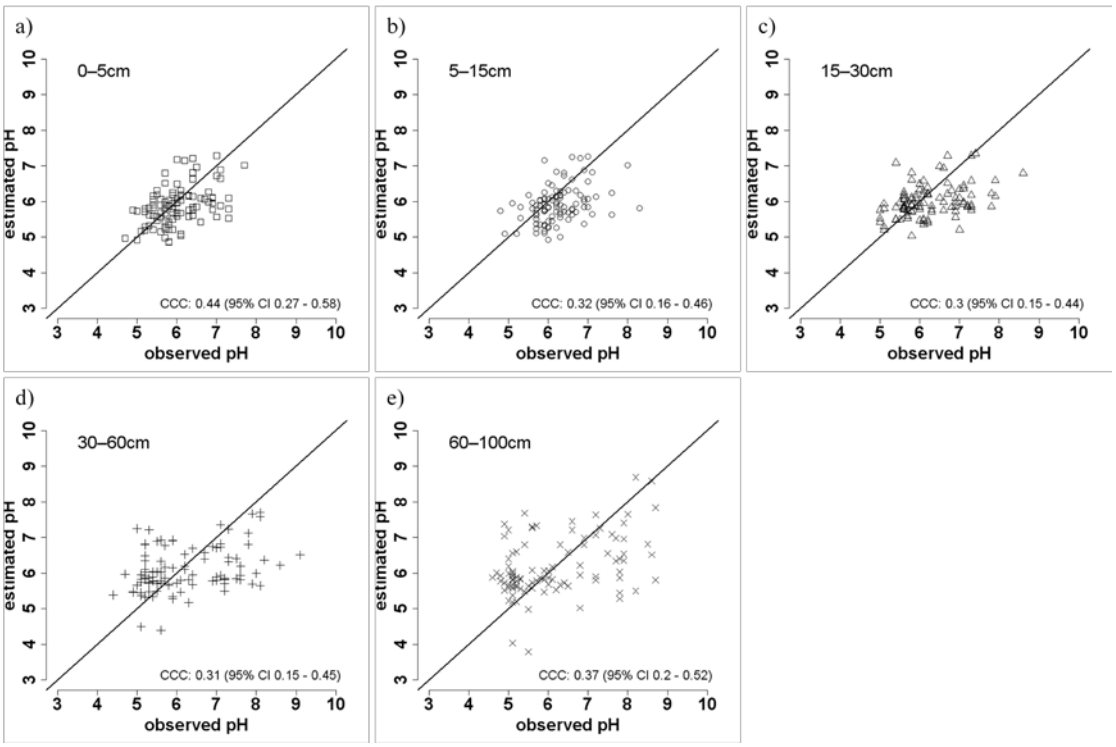


Figure 4.4.1. Plots of the observed soil pH vs. the corresponding digital soil mapped prediction of soil pH and resultant Lin's Concordance Correlation Coefficient at (a) 0-5 cm, (b) 5-15 cm, (c) 15-30 cm, (d) 30-60 cm, and (e) 60-100 cm.

What can also be observed from the plots is that at higher pHs (>7) there is a systematic under prediction, particular at 15–30 cm, 30–60 cm and 60–100 cm (Figure 4.4.1c-e). Bias estimates corroborate this observation approximately where strong positive bias (under prediction) was observed for both 5–15 cm (0.3) and 15–

30 cm (0.2± 0.1). For the depth increments of 30-60 cm and 60-100 cm, bias is smaller relative to the other depths; however the low accuracy can be attributed mainly to the higher level of imprecision at these sub-soil depths.

For further analysis, mainly in assessing the quality of the uncertainty estimation, Figure 4.4.2 shows the deviation of the areal proportion of the map correctly predicted at the corresponding confidence levels for each depth increment. The PICP plots demonstrate a significant degree of sensitivity with change in confidence level. Between the 90% and 40% confidence levels it can be observed that the areal proportions demonstrate a pattern of increasing deviation from the desired result. This outcome is likely attributed to the bias observed within the predictions which has resulted in the bias being transferred to the PIs because they are empirically derived from the prediction errors. At higher confidence levels, the PI performs as it would be expected to, and also has the added benefit of being able to buffer the bias in the predictions. For predictions that have significant bias, it is with decreasing confidence levels that it becomes evident that there is some mis-specification of the PIs at these confidence levels. The fact that the observed deviations are below the 1:1 line, indicates not only some misspecification of PIs, but more importantly is that the predictions of uncertainty are underestimated. In this regard this is better than the alternative where the areal proportions for a given confidence level are above the lines which would indicate the PIs are unnecessarily wide or in other words, an over-estimation of the uncertainty.

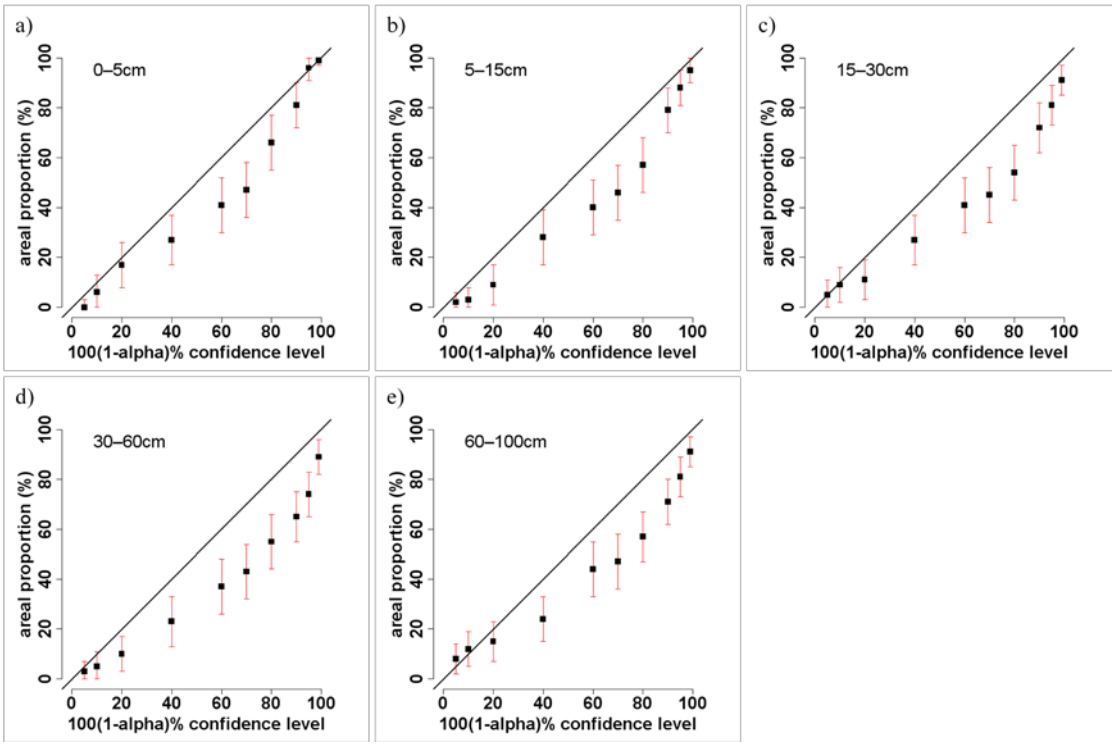


Figure 4.4.2. Prediction interval coverage probability plots for the areal proportion correctly predicted at (a) 0-5 cm, (b) 5-15 cm, (c) 15-30 cm, (d) 30-60 cm, (e) 60-100 cm.

As expressed in Chapter 3 the aim of calculating uncertainties is to account for all perceived or known sources, including those associated with our poor understanding of soil patterns and processes, and those associated with the model and their parameters and model inputs (covariates). Because the approach to quantifying the uncertainties is evaluated empirically, bias in the predictions will inadvertently mean bias will be present also in the estimates of uncertainty. This phenomenon was apparent in this study; the problem is that I used legacy soil data to generate the map which ultimately resulted in prediction bias. Consequently, this bias is reflected in the uncertainties as well. Because a probability sample was used to validate the pH map and their uncertainties I am able to discover such bias and subsequently an under-prediction of uncertainties. It is not always the case that independent data from a probability sample can be used for validation. Data splitting was used in Chapter 3, and at confidence levels from 5-99% there was a near matching proportion of observations which fitted within their PI for both available water and organic carbon. While a better result *per se*, it is clear to see that indicators of map quality are more valid when using an independent probability sample for validation.

From a map producer's perspective, there is significant value in coupling the proposed criteria with those conventionally reported for soil map quality. Firstly, the APCP, RMSES and RMSE all indicated an increased uncertainty with soil depth. I am more confident in the quality of the soil map at the soil surface where the predictions themselves are more accurate, but more importantly, the uncertainties of those predictions are also adequately accounted for. With increasing soil depth however, map quality decreases; there is decreased accuracy and precision of the actual predictions, coupled with a systematic underestimation of the uncertainties.

Overall (at all depths) the issues of bias and imprecision need to be addressed in order to improve the accuracy. Ideally this would be done directly by improvement in the modelling of the spatial distribution of the soil properties, which would result in a decrease of the RMSES. The PIs would naturally adjust themselves accordingly and become narrower in the process.

There are difficulties however with existing methods to make improved predictions within a DSM, particularly in the subsoil. One difficultly is that soil pH is evidently a dynamic soil property, which is likely to change as a result of human intervention such as agronomic practices (Bastida et al. 2008). From field knowledge of the study area, I expect however that significant change in pH as a result of intervention would only be minor. Rather, the difficulty is more the paucity of well known and available covariates that are able to describe soil attribute variability in the sub-surface.

A question that then remains is whether other soil properties can be validated using the same sampling units used in this study? In short, it is not optimal to do so. The procedures for validation in this study were optimised based on the predictions and their uncertainties of soil pH. Because of this, the stratification may be inappropriate for other soil attributes (maps) that could be validated in this area. The intention of this project was not to validate multiple soil attributes; it was more the presentation of additional criteria for quantifying map quality. Nevertheless, for the sake of efficiency there may be a requirement that multiple soil attribute maps need to be validated concurrently. One alternative to stratification on the basis of the predictions and their uncertainties is compact geographical stratification (Brus et al. 1999; Walvoort et al. 2010). With this design, the target area is stratified on the basis of spatial coordinates, resulting in strata being independent of the soil map. With such a design, the same design-based inference could be used as that used in this

study (de Gruijter et al. 2006). Future studies will obviously need to properly investigate this alternative approach.

4.5. Conclusions

In the course of this chapter, first I impressed upon the need to assess the quality of the quantifications of uncertainty in a DSM framework as one does for assessing the quality of the predictions alone. Subsequently, I presented two new criteria that collectively address map quality in terms of both the predictions *and* their uncertainties. These criteria are largely based on the empirical coverage of PIs- the methodology for expressing prediction uncertainty. The MSES explicitly deals with the prediction accuracy in that it is a modification of the MSE, yet includes a measure of the map uncertainty within its formulation. The APCP deals mainly with the quality of the uncertainty component whereby I express with 95% confidence that the true value of a soil attribute lies between the two interval limits. I used additional samples collected from a probability sample to determine unbiased estimates of these quality measures in addition to conventional quality measures such as MSE, ME and imprecision. Coupling these two new criteria with conventional measures means more information is gained. And for a map producer, aids in the efforts to improve precision of the map. For the map user, greater clarity of decision making regarding for example the optimisation of inputs (fertilisers etc), monitoring soil changes or ameliorating soil threats can be made. Regardless of the purported quality, it is up to the map user to determine the map's fitness for use. The criteria I have proposed in this study ensure a more objective approach to those decisions.

4.6. References

Bastida, F., Zsolnay, A., Hernandez, T., Garcia, C., 2008. Past, present and future of soil quality indices: A biological perspective. Geoderma 147: 159-171.

Bell, S.A.J., 2004. Vegetation of Werakata National Park, Hunter Valley, New South Wales. Cunninghamia 8: 331–347.

Bishop, T.F.A., McBratney, A.B., Whelan, B.M., 2001. Measuring the quality of digital soil maps using information criteria. Geoderma 103: 95-111.

Brown, J.D., Heuvelink, G.B.M., 2005. Assessing uncertainty propagation through physically based models of soil water flow solute transport. In: M. Anderson (Ed.), Encyclopaedia of Hydrological Sciences. John Wiley and Sons, Chichester, pp. 1181–1195.

Brus, D.J., de Gruijter, J.J., 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma 80: 1-44.

Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. European Journal of Soil Science 62: 394-407.

Brus, D.J., Spatjens, L., de Gruijter, J.J., 1999. A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. Geoderma 89: 129-148.

Burrough, P.A., Beckett, P.H.T., Jarvis, M.G., 1971. Relation between cost and utility in soil survey (I-III). Journal of Soil Science 22: 359–394.

Congalton, R.G., Green, K., 2009. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press, Boca Raton.

de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Knotters, M., 2006. Sampling for Natural Resource Monitoring. Springer-Verlag Berlin Heidelberg, The Netherlands.

Finke, P.A., 2007. Quality assessment of digital soil maps: producers and users perspectives. In: P. Lagacherie, A.B. McBratney, M. Voltz (Eds.), Digital Soil Mapping, an Introductory Perspective. Developments in Soil Science. Elsevier, Amsterdam, pp. 523–542.

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modelling approaches. Geoderma 152: 195-207.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference and Prediction. 2nd ed. Springer, New York, N.Y.

Isbell, R., 2002. The Australian Soil Classification. CSIRO Publishing, Collingwood, VIC.

Lark, R.M., 1995. Components of accuracy of maps with special reference to discriminant-analysis on remote sensor data. International Journal of Remote Sensing 16: 1461-1480.

McBratney, A.B., de Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science 43: 159–175.

Minasny, B., McBratney, A.B., 2002. Uncertainty analysis for pedotransfer functions. European Journal of Soil Science 53: 417–429.

Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences 32: 1378-1388.

Odgers, N.P., McBratney, A.B., Minasny, B., 2011. Bottom-up digital soil mapping. I. Soil layer classes. Geoderma 163: 38-44.

Odgers, N.P., McBratney, A.B., Minasny, B., 2008. Generation of kth-order random toposequences. Computers & Geosciences 34: 479-490.

Rayment, G.E., Higginson, F.R., 1992. Australian laboratory handbook of soil and water chemical methods. Australian soil and land survey handbook, vol 3. Inkata Press, Melbourne.

Shrestha, D.L., Solomatine, D.P., 2006. Machine learning approaches for estimation of prediction interval for the model output. Neural Networks 19: 225-235.

Solomatine, D.P., Shrestha, D.L., 2009. A novel method to estimate model uncertainty using machine learning techniques. Water Resources Research 45, Article Number: W00B11.

Thackway, R., Cresswell, I.D., 1995. An Interim Biogeographic Regionalisation for Australia: A Framework for Establishing the National System of Reserves, Version 4. Australian Nature Conservation Agency, Canberra.

Tranter, G., Minasny, B., McBratney, A.B., 2010. Estimating Pedotransfer Function Prediction Limits Using Fuzzy k-Means with Extragrades. Soil Science Society of America Journal 74: 1967-1975

Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Computers & Geosciences 36: 1261-1267.

Webster, R., Beckett, P.H.T., 1968. Quality and usefulness of soil maps. Nature 219: 680–682.

White, R.E., 2003. Soils for Fine Wines. Oxford University Press, New York.

Appendix 4.1. Accounting for measurement error in map validation

Measurements that are used for validation may be so accurate that one can safely assume that the effects of measurement error on the validation results are negligible. That assumption cannot be made in this study, thus we have to consider the effects of measurement error. The validation criteria used in this study are areal proportion correctly predicted, Root Mean Squared Simulation Error, Root Mean Squared Prediction error, Mean Error, and Imprecision. All of these, except for the Mean Error, involve non-linear transformation of the measured pH: a 0/1 indicator transformation for the APCP and squaring for RMSE and Imprecision. This implies that, without the bias corrections as detailed hereafter, the estimates of these 'non-linear' criteria would be biased. Especially the RMSE would be over-estimated, thus punishing the map for errors in the validation data, and all error variances would be underestimated.

Estimation of the areal proportion correctly predicted

As defined previously the criterion is defined as:

$$APCP = \frac{1}{\|A\|} \int_{s \in A} i(s) \mathrm{d}s$$

[a4.1.1]

where *A* is the mapped area, and *i(s)* equals 1 if the true pH value at location *s* is covered by the PI given by the map at *s*, and 0 otherwise. As we do not possess true pH values, we have no error-free *i's* either. Instead we have to work with indicator values determined from the measured pH values. These indicator values are thus subject to random error and denoted here by *I(s).* Each *I(s)* follows a Bernoulli distribution with expectation *π(s),* being the probability that a randomly measured pH value at *s* is covered by the PI at *s* (The pH measurements are assumed to have no systematic error.). The usual design-based estimator for *APCP* is:

$$\widehat{APCP} = \frac{1}{\|A\|} \sum_{h=1}^{H} \frac{\|A_h\|}{n_h} \sum_{j=1}^{n_h} I_{hj}$$

[a4.1.2]

where $\|A_h\|$ and n_h are the surface area and sample size of stratum *h*, respectively, and I_{hj} is the indicator value as determined at the *j-th* sample point of stratum *h*. If the pH measurements were error-free, *I(s)* would equal *i(s)* for all *s*, and $\widehat{APCP}$ would be an unbiased estimator. This is not so in the present study. To investigate the bias of $\widehat{APCP}$ we take its expectation over the process of measuring, conditional on a given sample $\boldsymbol{S}$:

$$E_m\{\widehat{APCP}|\boldsymbol{S}\} = \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h}E_m\{I_{hj}\}$$
$$= \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h}\pi_{hj}$$

[a4.1.3]

The conditional bias due to measurement error thus equals:

$$E_m\{\widehat{APCP}|\boldsymbol{S}\} - \widehat{APCP}_t = \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h}\{\pi_{hj} - i_{hj}\}$$

[a4.1.4]

where $\widehat{APCP}_t$ is the (hypothetical) estimate based on error-free measurements. We could assess this bias numerically by simulation, but in the present case we may assume that it is small enough to be neglected. The reason is that the terms $\pi_{hj} - i_{hj}$ tend to be small and, more importantly, being both positive and negative they will largely cancel out.

Estimation of the error variance of $\widehat{APCP}$

The error variance of $\widehat{APCP}$ due to sampling *and* measurement error is estimated by:

$$\hat{V}(ACPA) = \frac{1}{A^2}\sum_{h=1}^{H}\frac{A_h^2}{n_h(n_h-1)}\sum_{j=1}^{n_h}(i_{hj} - \overline{i_h})^2$$

[a4.1.5]

where $\overline{i_h}$ is the mean of the indicator values in stratum *h*.

Estimation of the Mean Square Error

The mean squared prediction error is defined as:

$$MSE = \frac{1}{\|A\|}\int_{s\in A}\{z_p(s) - z(s)\}^2 \mathrm{d}s$$

[a4.1.6]

where $z_p(s)$ and $z(s)$ are the predicted and true values at location *s* of the target variable, here pH. The MSE is estimated from stratified random sample data by:

$$\widehat{MSE} = \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h}(z_{phj} - z_{hj})^2$$

[a4.1.7]

Applying this estimator to measured values (z_m) instead of true ones gives:

$$\widehat{MSE} = \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h}(z_{phj} - z_{mhj})^2$$

[a4.1.8]

Assuming that the prediction error and the measurement error are spatially uncorrelated, this can be re-written as:

$$\widehat{MSE} = \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h}\{(z_{phj} - z_{hj})^2 + (z_{mhj} - z_{hj})^2\}$$

[a4.1.9]

Taking the expectation over both sampling and measuring gives:

$$\begin{aligned} E_p E_m \widehat{MSE} &= MSE + \frac{1}{\|A\|}\sum_{h=1}^{H}\frac{\|A_h\|}{n_h}\sum_{j=1}^{n_h} E_p E_m\left(z_{mhj} - z_{hj}\right)^2 \\ &= MSE + \sigma_m^2 \end{aligned}$$

[a4.1.10]

where σ_m^2 is the variance of the measurement error. It follows from (a4.1.10) that the usual estimate (a4.1.8) should be diminished with σ_m^2 to make it unbiased.

Estimation of the Error Variance of $\widehat{MSE}$

The error variance of $\widehat{MSE}$ due to sampling and measurement error is estimated by:

$$\hat{V}(MSE) = \frac{1}{A^2}\sum_{h=1}^{H}\frac{A_h^2}{n_h(n_h-1)}\sum_{j=1}^{n_h}(d_{hj}^2 - \bar{d}_h^2)^2$$

[a4.1.11]

where d_{hj}^2 is the squared difference between the predicted and the measured value at location *j* in stratum *h*, and $\bar{d}_h^2$ is the stratum mean of the squared differences.


support
measurement
concentration
upscaling
farm
model
resolution
spacing
digital
Soil
available
target
one
different
pixel
variable
may
methods
areal
problem
situations
kriging
information
points
operations
involves
problems
spatial
area
soil
processes
downscaling
scales
process
outputs
convolution
uncertainty
data predictions
estimates
procedure
extent
errors
thorium
coarse-gridding
point
block
deconvolution
example
grid
used
observations
disseveration
error ppm
described
scaling
equal
value
required
prediction
changes mapping
cell
changing
involve
maps
Upscaling
scale
map

Customer *(walks into shop or milk bar and addresses shopkeeper): Bottle of milk thanks.*

Shopkeeper: *(answering the customer in a frowning sort of way) Low fat, no fat, high fat, full cream, high calcium, high protein, soy, lite, skim, omega-3, high calcium with vitamin D and folate , or extra dollop?*

(Pregnant pause)

Customer: *Ah (quiet chuckle) I just want milk that tastes like real milk.*

.....

[Australian television advertisement (2010)]

Chapter 5

Snakes and ladders: manipulations of scale for digital soil mapping

Summary

Part of making informed decisions regarding land management in an agronomical or ecological setting often requires having available, spatially explicit soil information. Whether the soil information that is available is compatible with what is required in terms of spatial scale is often questionable. Rather than continually conducting new soil survey to complement decision-making, one way to deal with the incompatibility issue is via manipulations of scale on the existing soil information. Manipulations of scale for digital soil mapping involve changes in map extent, grid cell resolution and prediction support. Subsequently, this chapter introduces explicit terminology to describe different forms of scale manipulation in the context of digital soil mapping. Furthermore, a suite of existing pedometric methods is described for implementation of each process. The different forms of scale manipulation are described in terms of changes to grid cell spacing and prediction support. Fine-gridding and coarse-gridding are operations where the grid spacing changes but support remains unchanged. Deconvolution and convolution are operations where the support always changes which may or may not involve changing the grid spacing. While disseveration and conflation operations occur when the support and grid size are equal and both are then changed equally and simultaneously.

5.1. Introduction

The previous three chapters investigated practicable methodologies for digital soil mapping concerning whole-profile predictions, quantifying prediction uncertainties, and validating the predictions *and* their uncertainties. For the next two chapters, the focus shifts to matters of scale and manipulations thereof.

Digital soil mapping involves the combination of soil observation (from field soil survey) and numerical methods to populate continuous spatially explicit soil information databases; the digital soil map being the final visual communicative product (Lagacherie 2008). Digital soil maps and associated soil information databases are used to enable land managers, modellers and policy makers to better understand and manage the soil resource either for agricultural benefit or for some other ecological purpose.

The operational status of digital soil mapping now means that the availability of digital soil information products encompasses a hierarchy of spatial scales which include global, continental, national, region, farm and field scales (Grunwald et al. 2011). Soil information availability may not be an issue *per se*; however, an often encountered problem is it is incompatible to meet the objectives of a given project or policy directive. The incompatibility issue is largely a matter of scale dissimilarity; soil information may be available at one scale, but may be required either at a finer or coarser scale and maybe even required at a different support or volume (Papritz et al. 2005). For example, digital soil maps created from point support measurements (soil cores, pits etc.) will generate point support maps which may not be of any use when a policy directive requires the support of predictions to be blocks or a specified land unit size. Rather than enduring with what is available, a way to get more value out of existing soil maps is to implement either upscaling or downscaling methods with the intention of creating soil information products that are nuanced with the requirements of the end users i.e. the land managers, modellers and policy makers.

Downscaling and upscaling of soil information has previously received considerable attention. Various soil science workshops have been conducted, e.g. Finke et al. (1998) detail the breadth of issues with many examples concerning scale in the soil and water sciences domain. McBratney (1998) made some suggestions for a number of possible approaches for upscaling or downscaling soil information problems. Similarly, Bierkens et al. (2000) developed and presented a general

framework in the form of a decision tree to detail processes and their models for solving various scaling problems. Issues of scale incompatibility are not unique to the soil science domain, however. In climatology research, outputs of climate simulations from general circulation models (GCMs) cannot be directly used for hydrological impact studies of climate change because of a scale mismatch (Bloschl 2005). The grid resolution of GCMs is generally in the order of hundreds of kilometres. In contrast, the resolution at which inputs to hydrological impact models are needed is in the order of tens or hundreds of square kilometres. Practitioners in the remote sensing domain also, in order to understand the underlying geophysical process of some atmospheric or environmental variables, often use two or more instruments which measure the same processes, but measure it at different spatial supports (Nguyen et al. 2010). As a consequence scaling methods are required to combine both information sources for making optimal inferences of the underlying process. Scaling, as an operative process, is essentially an inference of spatial processes at one resolution from data at another resolution; which in spatial statistics is often called the "change-of-support" problem (Cressie and Wikle 2011). The change-of-support problem presents many statistical challenges and has been reviewed in Gotway and Young (2002) with other important contributions from Cressie (1996) and Fuentes and Raftery (2005).

The aim of this chapter is to describe a non-exhaustive suite of existing pedometric techniques or processes for conducting scale manipulation of digital soil maps that take into account the extent, resolution and support of the source (map to be scaled) and destination (new map) maps.

5.2. Some concepts

5.2.1 The digital soil map model

The raster model is the principal format on which digital soil maps are displayed, where each pixel or grid cell (the single unit entity of a raster), which has a spatially explicit location, contains a value for a given target soil property. Digital soil maps have three scale entities; extent, resolution, and support for which Western and Bloschl (2005) termed the scaling triplet when discussing scaling issues for hydrological modelling. Map extent is the areal expanse or coverage, such that the map could be a soil map of the world, a country, a region, or a particular farm.

Resolution is the grid cell spacing or pixel size of the raster. A map made up of pixels which have dimensions of 10m × 10m is a map with a resolution of 10m. While support is likened to a volume or area which the predictions are representative of; this could either be points (that have no defined area or volume) or blocks (which have a measurable area or volume). The scaling triplet is conceptualised by a generic soil map model (Bishop et al. 2001) which consists of a soil variable, which is estimated with some uncertainty. This variable is predicted onto grid cell spacing, *G* that has a support *B* which could be a point or a block (Figure 5.2.1).

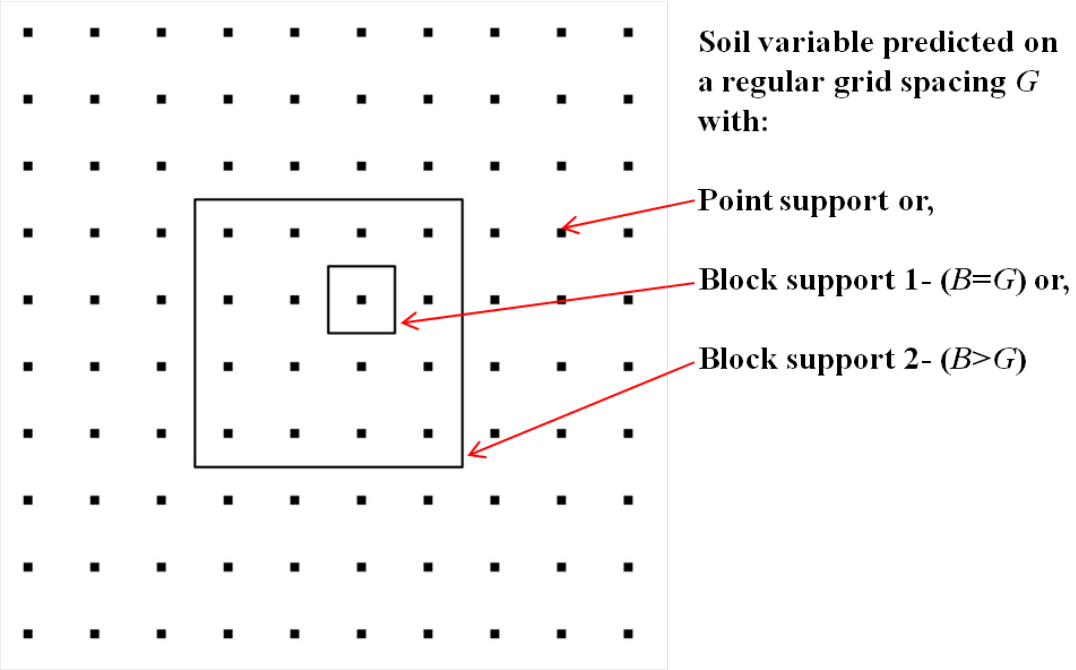


Figure 5.2.1. Generic soil map model (adapted from Bishop et al. 2001). Support of predictions are point when block *B* is very small. Block support prediction occur when *B* is greater than 0. Block support 1 is when *B* equal grid spacing *G*. Block support 2 is when *B* is greater than *G*. *B* may be larger than grid spacing.

The soil map model is the same as a raster model when *G* is equal to *B* (block support 1 of Figure 5.2.1). Because *B* has some definable dimensions or support it has some area or volume and the value attributed to it represents an averaged value for that area or volume. When *B* is very small, the map model is essentially a grid of points; the support in this case is a point (point support of Figure 5.2.1). Despite both maps being displayed on the same raster model because they have equal *G*, they are fundamentally different because they have different supports; the values attributed to the pixels mean different things. For further complexity, *B* may even be bigger than

G, which is quite common in situations where block kriging is used (block support 2 of Figure 5.2.1). A situation where this would be used, as discussed by Bishop et al (2001), is where a map producer using block kriging may want a dense coverage of information (finely spaced *G*), but to reduce the uncertainty of prediction may choose a much larger *B* than *G*.

5.2.2. Scaling digital soil maps

This chapter addresses the posterior problem of how does one without additional sampling, perform either upscaling or downscaling of a map of some target variable across a particular extent with a resolution *G* and support *B.*

Manipulations of the map extent can and usually are coupled with increasing or decreasing the grid cell spacing in order to upscale or downscale respectively (McBratney 1998). However in accordance with the digital soil map model where predictions have some sort of spatial support, upscaling and downscaling may be too general or nebulous in meaning. Aggregation and disaggregation have an equivalent meaning to upscaling and downscaling in soil science (Bierkens et al. 2000). However these terms are also used frequently to describe procedures for combining or separating traditional soil map class/units respectively. To this end, new terminology needs to be introduced that provides explicit meaning to scaling methods for digital soil maps both in terms of moving up or down the hierarchy of scales and through increasing or decreasing the support of the predictions.

The four panels of Figure 5.2.2 illustrate contrived examples of digital soil maps, all of which have the same spatial extent. Panel 1 (P1) and Panel 2 (P2) are the same raster model, but different digital soil map models. The grid spacing is the same but in P1 the support is a point, while in P2 the support is a block where *B* has dimensions equal to *G.* The situation is the same when comparing P3 with P4. Obviously P1 has finer grid spacing than P3 and it should be assumed that in the hierarchy of scales P1 is below P3 meaning a smaller scale and in reality may or may not have a smaller extent. It is also to be assumed that P1 and P2 exist on the same level of the scale hierarchy.

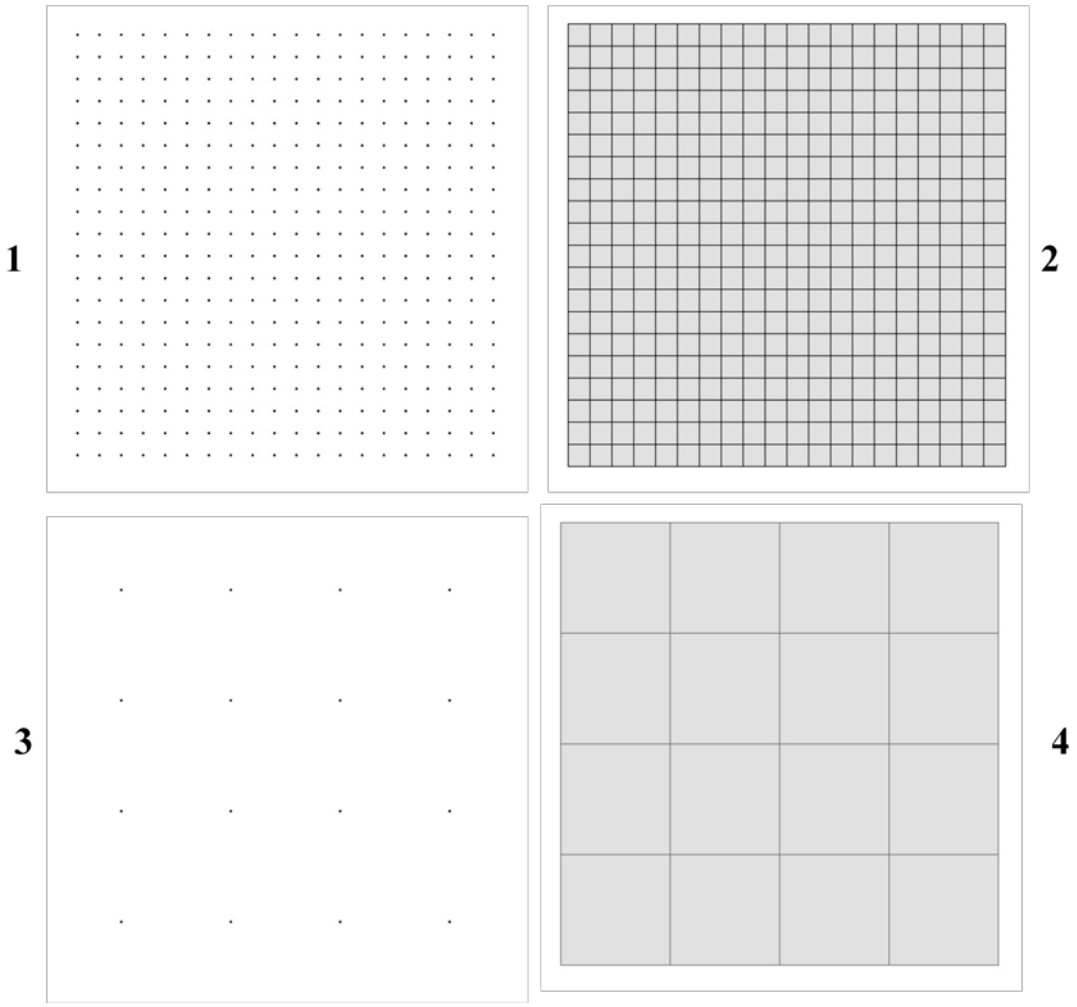


Figure 5.2.2. Exemplar soil map models. Panel 1 and Panel 2 have the same grid spacing yet Panel 2 is on block support (where block size is equal to the grid spacing), Panel 1 is on point support. Similarly for Panel 3 and Panel 4 except the grid spacing is larger.

The intent of Figure 5.2.2 is to provide some visual reference for scaling methods with regards to the previously defined digital soil map model. The scaling methods are summarised in Table 5.2.1 and is to be interpreted by deciding firstly which soil map model suits the source map with a corresponding row selection. This is followed by a column selection of the digital soil map model that is the desired scale destination of the new map. The row and column coordinate pair then refer to the scaling method required to perform the process. The scaling methods are:

Fine-gridding and Coarse-gridding:- These are situations where *G* changes but *B* remains unchanged. Examples of fine-gridding are situations where a process of moving from Panel 3 to Panel 1 (Figure 5.2.2; P3→P1) is required. While coarse-gridding situations involve P1→P3 processes.

Deconvolution and Convolution:- These are situations where *B* always changes which may or may not involve changing *G*. However when both *B* and *G* are equal and changed simultaneously, the changes *are not* equally applied. Convolution processes always involve an increase in *B*, with the examples being P1→P2, P1→P4, P3→P2, P3→P4 operations. Deconvolution always involves a decrease in *B* such as P2→P1, P2→P3, P4→P1 or P4→P3 processes.

Disseveration and Conflation:- These are situations when *B* and *G* are equal and both are changed equally and simultaneously. Examples of disseveration are situations where a P4→P2 process is required. Conflation situations involve P2→P4 processes.

Table 5.2.1. Coordinate table of scaling processes based on attributes of the source map (rows) and scale attributes of destination map (columns).

	Process			
	Points (fine)	**Blocks (fine)**	**Points (coarse)**	**Blocks (coarse)**
Points (fine)		Convolution	Coarse-gridding	Convolution
Blocks (fine)	Deconvolution		Deconvolution	Conflation
Points (coarse)	Fine-gridding	Convolution		Convolution
Blocks (coarse)	Deconvolution	Disseveration	Deconvolution	

5.3. Specific discussion of scaling methods for digital soil mapping.

This section explores in more detail some pedometric procedures for performing each of the described scaling processes. When considering the hierarchy of spatial scales recognised for soil (Hoosebeek and Bryant 1992), the *i*-levels of interest for applying the described scaling processes in the context digital soil mapping would be global (*i*+6), continental (*i*+5), region (*i*+4), watershed (*i*+3), farm (*i*+2) and field (*i*+1) scales. The methods discussed are not exhaustive and focus on one or two procedures that could be implemented to create the outputs at the desired scale.

5.3.1. Fine-gridding and coarse-gridding

The most common scaling methods that would be encountered in digital soil mapping would be either fine-gridding or coarse-gridding operations. These are operations where the grid spacing *G* changes without any change of the support i.e. *B* remains constant. Considered point-to-point processes, fine-gridding is a downscaling problem whereas coarse-gridding is an upscaling problem.

Fine-gridding requires some form of point interpolation or spatial prediction. A stochastic process such as ordinary punctual kriging may be used to interpolate onto the finely resolved grid nodes. See Isaaks and Srivastava (1989) for more theoretical details of ordinary kriging. Assuming the mean is unknown; the values at the

interpolated point locations (fine resolution points) are treated as random variables and are estimated from surrounding point predictions at the coarser scale. The ordinary punctual kriging predictor is:

$$\hat{Z}(x_0) = \sum_{i=1}^{N} \lambda_i \cdot z(x_i)$$

[5.3.1]

where $\hat{Z}(x_0)$ is the value of the target variable at unvisited location x_0 which is predicted from a weighted linear combination of N neighbouring point observations $z(x_i)$ at the coarser scale with weights λ_i. To ensure an unbiased estimate the weights from the vector $\boldsymbol{\lambda}$ are made to sum to 1 and are obtained by solving the ordinary kriging system:

$$\begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{bmatrix} = \mathbf{A}^{-1} \cdot \mathbf{s}$$

[5.3.2]

where $\boldsymbol{\mu}$ is a Lagrange multiplier necessary to solve the system, **A** is a matrix with semi-variances between the data points at the coarse resolution and has the structure:

$$\boldsymbol{A} = \begin{bmatrix} \gamma(x_1, x_1) & \gamma(x_1, x_2) & \cdots & \gamma(x_1, x_N) & 1 \\ \gamma(x_2, x_1) & \gamma(x_2, x_2) & \cdots & \gamma(x_2, x_N) & 1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \gamma(x_N, x_1) & \gamma(x_N, x_2) & \cdots & \gamma(x_N, x_N) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}$$

[5.3.3]

where γ is semi-variance. The vector **s** contains the semi-variances between the coarse-scaled points and the fine-scaled point x_0 and has the structure:

$$\boldsymbol{s} = \begin{bmatrix} \gamma(x_1, x_0) \\ \gamma(x_2, x_0) \\ \vdots \\ \gamma(x_N, x_0) \\ 1 \end{bmatrix}$$

[5.3.4]

The kriging variance is equated as:

$$\hat{\sigma}^2(x_0) = \boldsymbol{s}^{\boldsymbol{T}}\boldsymbol{\lambda}$$

[5.3.5]

For DSM, the value of $z(x_i)$ will often be uncertain, such that the exact value will not be known. These uncertainties could be measurement or prediction errors and need to be accounted for in the kriging system. Kriging with uncertain data was introduced by Delhomme (1978) and the method requires some modification of the standard ordinary kriging equations. However, to do this we need to assume 1) the errors have a zero mean; 2) the errors are uncorrelated; 3) the errors are not correlated with the target variable; 4) the variance of the errors is a known quantity and varies from point to point (Delhomme 1978). Under these assumptions, following the formulations from Christensen (2011) the semi-variance elements (*i*, *j*) of the matrix **A** can be modified on the off-diagonals i.e. where $i \neq j$ to:

$$(i,j) \text{ element of } \mathbf{A} = \gamma\left(x_i, x_j\right) + \frac{\sigma^2(x_i) + \sigma^2\left(x_j\right)}{2} \quad \textbf{[5.3.6]}$$

where σ^2 is the measurement error variance. Furthermore, the *i*th element of the **s** matrix is modified to:

$$ith \text{ of element of } \mathbf{s} = \gamma(x_i, x_0) + \frac{\sigma^2(x_i)}{2} \quad \textbf{[5.3.7]}$$

Bear in mind these adjustments are only valid when the target variable is not correlated with the error variances. While more problematic, Christensen (2011) formulates a method for dealing with correlated error variances for kriging with uncertain data which is based upon using variance-stabilising transformations as proposed by Box and Cox (1964).

Before any of this can be done however, and to be consistent, one needs to correct the variogram for the measurement errors. A simple way to correct for the bias of the variogram due to measurement errors is to calculate the spatial average of the error variances $\left(\frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i)\right)$ and to subtract this average from the variogram (Christensen 2011). In practice, a semi-variogram is fitted to all $z(x_i)$ from which the variogram parameters of the nugget, partial sill and range, denoted as c_Z, v_Z,and r_Z respectively, are obtained. To correct the variogram we simply subtract $\frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i)$ from c_Z. In general the range and partial sill are unaffected by the bias correction (Christensen 2011). Occasionally the estimated nugget may be less than

the averaged measurement error variance. When the spatial average of the error variances subtracted from the nugget variance is less than zero (i.e. c_Z - $\frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i) < 0$), the nugget for the adjusted semi-variogram can be set to zero. The adjusted partial sill can be set to $v_Z + c_Z - \frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i)$, so that the sill (sum of the nugget plus partial sill) is still reduced by $\frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i)$ (Christensen 2011).

Continuing on with another method of fine-gridding; if fine scaled environmental covariate data is available, a deterministic empirical or data mining approach could be implemented for which an overview of methods is given in McBratney et al (2000). Combining both deterministic and stochastic processes through a regression kriging approach is also another viable option; the method of which is detailed by Odeh et al. (1995). There is a good logical consistency in transfer between the hierarchies of scales using available covariate information; we know that the variation of soil properties depends on factors such as parent material, climate, land use and topography. These factors all operate at different scales and therefore influence soil processes and soil variation at different scales (Addiscott 1993). For example, a soil map of carbon variability may be available at 1km resolution over a large extent for which one could discern that climatic variables maybe be dominant factors controlling the carbon distribution. One then may want to use this map to determine what the carbon variability is at a particular farm at a resolution more interpretable for farm management decisions; for example, at 25m resolution. Climatic variables may be inconsequential at the scale of a farm, rather factors such as landuse or landscape position may be better able to explain the soil carbon variation. Thus, when using covariate information for fine-gridding operations it is important to consider what factors are dominating at the scale for which information is being downscaled to. This information is likely to be garnered from expert knowledge of the target variable and how it varies across different scales.

Coarse-gridding is a trivial upscaling exercise and is popularly practiced within Geographical Information Science (GIS) environments through such operations as re-sampling fine-gridded data to a coarser resolution. Nearest-neighbour samplings in addition to averaging and smoothing spline type operations are popular re-sampling methods. One must be careful that in the context of coarse-gridding, *B*

remains constant in the scaling procedure. Therefore in the context of the digital soil map model, regardless of whether an averaging or smoothing spline re-sampling procedure is used, the upscaled soil information product will still be on point support. A crude way of describing what coarse-gridding does, is that the map producer is effectively throwing some data away. The purpose of this may be because of a computer memory saving reason or that a particular map at a fine-scale is difficult to interpret and by performing coarse-gridding the map becomes simpler.

5.3.2. Deconvolution and convolution

Manipulations of scale that involve changing the support coupled with or without changing the grid spacing involve either deconvolution or convolution. Convolution is an upscaling problem because all situations entail increasing the support of the predictions, for example, point-to-block operations. Deconvolution is a downscaling problem where always the support size is decreased i.e. a block-to-point operations. For both convolution and deconvolution, changing the support is always performed, but changing the grid spacing is not always necessary.

5.3.2.1. Convolution problems

One form of convolution includes P1→P4 scaling processes. Here the grid spacing increases in addition to an increase in the support of the predictions. Because each block has many point observations, convolution could involve averaging the point observations contained within each block or pixel (Bierkens et al. 2000), such that:

$$\bar{Z}_H = \frac{1}{N}\sum_{i=1}^{N} z(x_i)$$

[5.3.8]

where the prediction of a single block Z with support H is obtained as an average of all $z(x_i)$ within H. The variance is then computed as:

$$\hat{\sigma}^2(\bar{Z}_H) = \frac{1}{N-1}\sum_{i=1}^{N}[z(x_i) - \bar{Z}_H]^2$$

[5.3.9]

It is necessary to indicate the variance so as to derive a confidence interval about the block average because the estimate is based only on a limited (not exhaustive) number of points. However the derivation of a confidence interval is based on the assumption that the *N* points are independent and that the sample mean follows a normal distribution. Brus and de Gruijter (1997) state that independence can be created through randomisation of the point locations. For P1→P4 processes the distribution of the points will not be randomly distributed; they will in fact be regularly spaced points. This means that the coverage of points within each block may be useful in the practicable sense of deriving a meaningful block average, yet the suitability of this method from a statistical view is not optimal. While not optimal, the suitability of implementing this particular P1→P4 process will rely on having many (for example >50) points within each block.

Alternatively when there are a sparse number of points inside each block, ordinary block kriging could be used (Burgess and Webster 1980). This geostatistical method is suitable because the aim of block kriging is to predict the mean value of a target variable in a region *V* of area *H* that centres on a point at x_0. The block kriging estimator is defined as:

$$\hat{Z}_H(x_0) = \sum_{i=1}^{N} \lambda_i \cdot z(x_i)$$

[5.3.10]

The predictor *Z* with support *H* is obtained from a weighted linear combination of *N* neighbouring point observations $z(x_i)$. The weights (λ_i) are obtained by solving the block kriging system which is the same as that for ordinary punctual kriging. In these cases the area of *H* will be the dimensions of the pixels to which the fine-scaled point observations need to be upscaled to. However, one difference between point and block kriging is the nature of the **s** vector, such that:

$$\boldsymbol{s} = \begin{bmatrix} \bar{\gamma}(x_1, x_0) \\ \bar{\gamma}(x_2, x_0) \\ \vdots \\ \bar{\gamma}(x_N, x_0) \\ 1 \end{bmatrix}$$

[5.3.11]

where $\bar{\gamma}$ is the average semi-variance between x_i and x_0 which is given by the integral:

$$\bar{\gamma}(x_i, H) = \frac{1}{|H|}\int_H \gamma(x_i, x)\,\mathrm{dx}$$

[5.3.12]

where $\gamma(x_i, x)$ denotes the semi-variance between the point x_i and a point x describing the block. The block kriging variance is equated as:

$$\hat{\sigma}^2(x_0) = \boldsymbol{s}^{\boldsymbol{T}}\boldsymbol{\lambda} - \bar{\gamma}(H, H)$$

[5.3.13]

where $\bar{\gamma}(H, H)$ is the within-block averaged semi-variance value.

Other convolution problems are those processes requiring scaling from P1→P2 or P3→P4. Consider the situation where a digital soil map may be available at point support where each pixel value represents a single point within the areal extent of the pixel (usually the central node). Without additional sampling, it may be necessary to know what the average of the target variable is across the entire area of each pixel. To perform this task, it is strictly a procedure where change of support is required without changing the grid size and can be performed using block kriging. To increase the support of a point map all that is required is to set the block size *H* equal to the grid spacing. An example of this is detailed in the following section.

5.3.2.2. Block Kriging example of P1→P2 processes

The gamma radiometric signal of thorium was collected using a proximal sensing device across fields at the I.A. Watson Grains Research Institute (30.27°S 149.80°E), a 460-ha property near Narrabri, in north-western New South Wales. Much of the property is under agricultural management where cereals (wheat, rye and triticale) have been bred under traditionally managed, irrigated cropping operations for more than half a century. For further details regarding this site, see Miklos et al. (2010). In the Summer of 2010, observation of thorium concentration (ppm) was made onto a regular grid of points with 5m spacing (point support map). Each independent observation also had some quantitative value of the measurement error given as a variance. Both maps are shown in Figure 5.3.1.

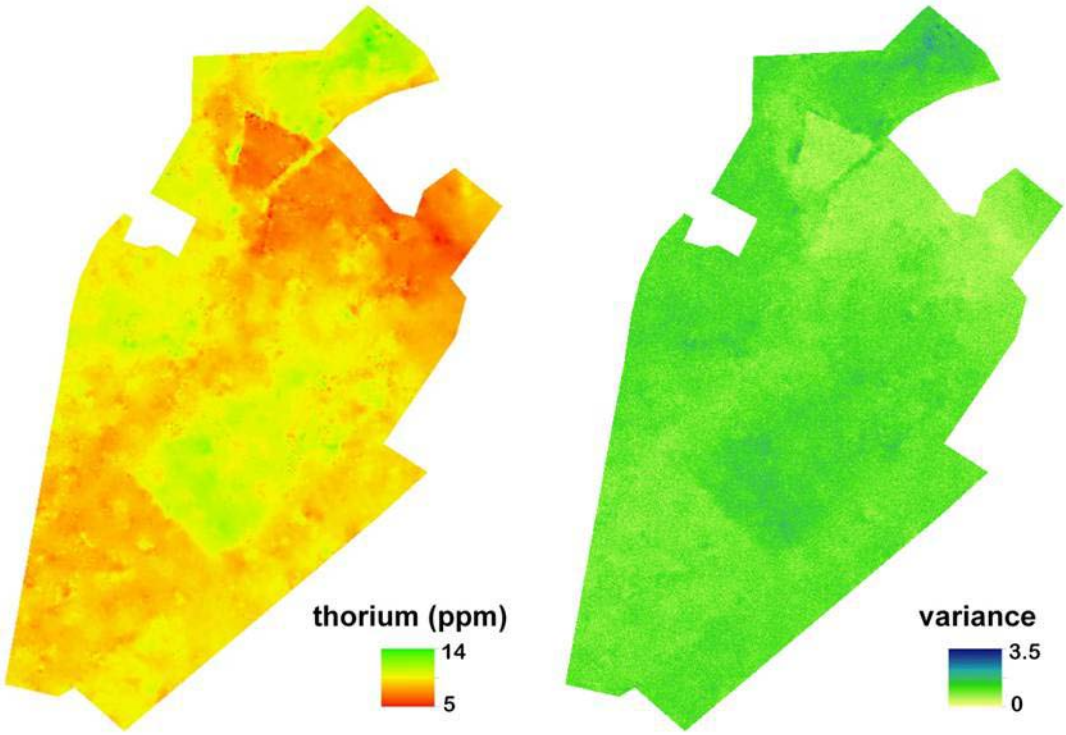


Figure 5.3.1. Panel 1- 5m point support thorium concentration (ppm) map. Panel 2- Associated variances of the measurement errors (also at 5m point support) of thorium concentration (ppm).

The aim of the example is to create new maps on block support with resolutions of 20m, 50m, and 80m, meaning that the block sizes are 20m×20m, 50m×50m and 80m×80m respectively. The reason why three increasingly larger resolutions are used is for comparative purposes in assessing the quality of the outputs of the P1→P2 processes.

Using the 5m point map, a simple way to generate block support maps at these desired resolutions and supports is to average all the observations within each block. This particular procedure is in fact a P1→P4 process (section 5.3.2.1). Because there is some uncertainty about the 5m spaced observations, we may arithmetically determine the average estimate of thorium (ppm) in each block as:

$$\overline{TB} = \frac{\sum_{i=1}^{N}\left(\frac{Tb_i}{s_i^2}\right)}{\sum_{i=1}^{N}\left(\frac{1}{s_i^2}\right)}$$

[5.3.14]

where $\overline{TB}$ is the weighted averaged value of a block, N is the number of point observations of the target variable Tb_i within each block, and with measurement error variance s_i^2. The variance of $\overline{TB}$ can be estimated by:

$$var(\overline{TB}) = \frac{1}{\sum_{i=1}^{N} \frac{1}{s_i^2}} \times \frac{1}{(N-1)} \sum_{i=1}^{N} \frac{(Tb_i - \overline{TB})^2}{s_i^2}$$

[5.3.15]

Figure 5.3.2 shows the block support maps in panel 1a, 1b and 1c. The spatial average of the $var(\overline{TB})$ was 2×10^{-3} , 7×10^{-4}, and 5×10^{-4} for the 20m, 50m, and 80m maps respectively. We could consider these maps as the 'true' block support maps and use them to compare with the outputs of the P1→P2 process, which is described now.

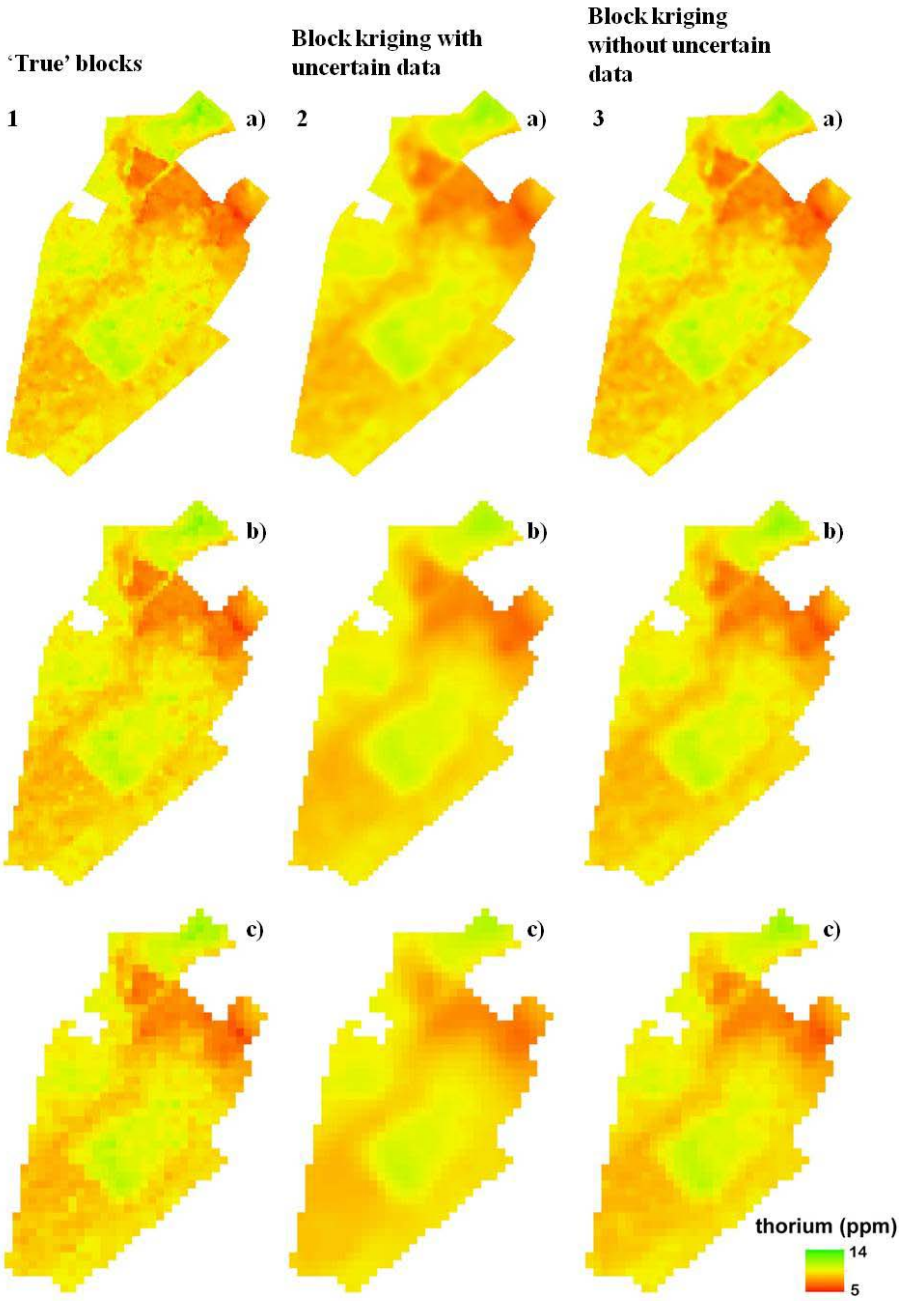


Figure 5.3.2. Block support maps of thorium (ppm) where support size equals grid cell size (resolution) - a) 20m, b) 50m, and c) 80m. Panel 1- 'true' blocks created directly from 5m point support map with a P1→P4 process- weighted averaging. Panel 2- P1→P2 process- Block kriging with uncertain data. Panel 3- P1→P2 process- Block kriging without including uncertainties.

Firstly, using the 5m point support map, coarse-gridding (P1→P3) was performed to generate new maps. For example, to create the 20m point support maps (thorium ppm and variance) we sample the 5m maps at grid nodes every 20m apart and so on. Block kriging is used (as described in section 5.3.2.1) to create the desired block support maps from the given point support maps where the support *H* is set to the

same size as the map resolution. However, because there is uncertainty about the true values of all Tb_i expressed as prediction or measurement error variances, there is a need to modify the standard ordinary block kriging equations. We can assume the error variances are independent of Tb_i and carry out what was described for punctual kriging with uncertain data in section 5.3.1 by modifying the **A** and **s** matrices accordingly which are formulated in Equations 5.3.3 and 5.3.4 respectively. Similarly the variogram of Tb_i is adjusted to correct for the bias due to measurement errors where the spatial average of the error variances $\left(\frac{1}{N}\sum_{i=1}^{N}\sigma^2(x_i)\right)$ is subtracted from the variogram (Christensen 2011). Because block kriging is being used, the *i*th elements of the **s** matrix is modified to:

$$ith \text{ of element of } \mathbf{s} = \bar{\gamma}(x_i, x_0) + \frac{\sigma^2(x_i)}{2} \qquad \textbf{[5.3.16]}$$

The block kriging variance is equated as in Equation 5.3.13.

Panel 2 of Figure 5.3.2 shows the maps that resulted from block kriging with uncertain data for each of the three resolutions and supports. For a comparative exercise, block kriging without including the uncertainties using the standard ordinary block kriging equations was also performed and the maps are shown in Panel 3 of Figure 5.3.2. While quite similar to the 'true' block support maps, including the measurement error variances into the kriging equations resulted in smoother representations of thorium concentration (ppm) at each of the three supports. The spatial average of the kriging variances from kriging with the uncertain data were 3×10^{-2}, 6×10^{-2}, and 9×10^{-2} for the 20m, 50m, and 80m maps respectively. The spatial averages of the kriging variances when not including the error variances was 2×10^{-3}, 3×10^{-2}, and 5×10^{-2} for the 20m, 50m, and 80m maps respectively. Essentially what these results represent is that for P1→P2 processes, the uncertainty increases with increasing resolution and support size. Logical also is the fact that the uncertainties are higher when kriging is performed using uncertain data compared with when the data is assumed to be without error. In any case the spatial averages of the kriging variances from both methods were higher than found for the 'true' blocks, which is to be expected.

The plots on Figure 5.3.3 illustrate the similarity of the maps resulting from kriging with uncertain data with the 'true' block maps. With 20m blocks, Lin's Concordance (CCC; section 2.2.5) was quantified as 0.97 while the R^2 was found to be 0.94. This indicates a high degree of similarity between the 'true' and predicted map. By not including the measurement error variances however, both the 'true' and predicted maps are very close to identical where a CCC of 1 was quantified. Similarly with the 50m blocks, the map resulting from kriging with uncertain data is a very good representation of the 'true' block map (CCC=0.95). With 80m blocks the CCC was found to be 0.93. At these two supports (50m and 80m), kriging without uncertain data resulted in outputs very similar to that of the 'true' blocks and better than that for when the measurement error variances were included.

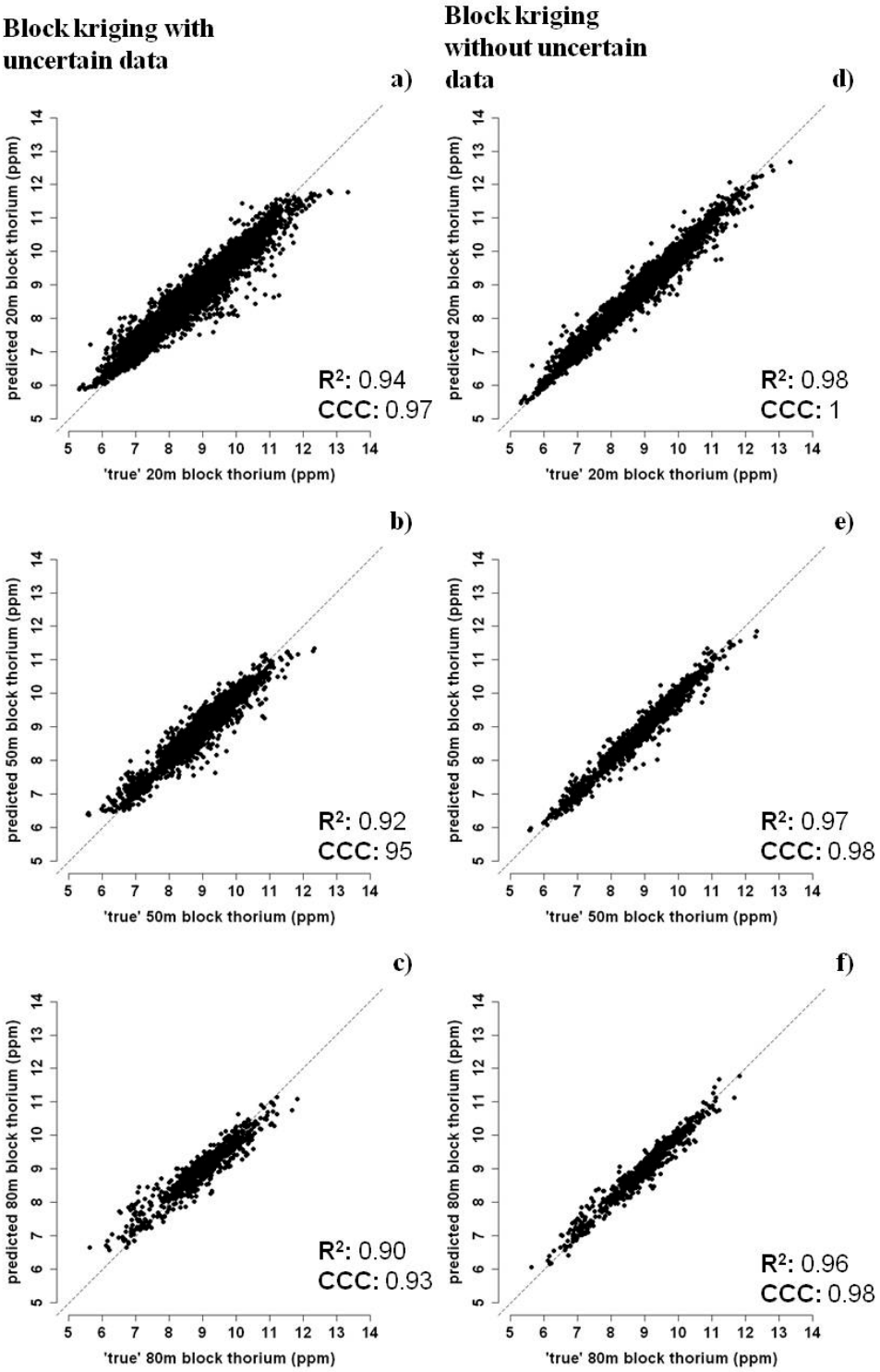


Figure 5.3.3. Map comparisons. Comparisons between 'true' block maps with maps from block kriging with uncertain data a) 20m, b) 50m, and c) 80m. Comparisons between 'true' block maps with maps from block kriging without uncertain data d) 20m, e) 50m, and f) 80m.

From this example of a P1→P2 process, block kriging tends to work better when smaller supports and resolutions are used. This is because more data close to the

location where a prediction is to be made is available. When uncertain data are used for kriging, the resulting maps will be smoother than when they are not. Empirically from this example, this is because more weighting (from the kriging weights) is assigned to points further away from the location where a prediction is to be made.

5.3.2.3. Further convolution problems

Convolution problems could also involve situations where one requires a process for scaling from P3→P2. The purpose for these processes may be that in addition to requiring point predictions to be expressed on an areal support, the target variable information is needed at a finer resolution to what is currently available. It is possible to achieve this directly through such methods as ordinary block kriging or universal block kriging. Because there is a need to describe the variation of a target variable at a finer resolution, ordinary block kriging would suit in situations where no available covariate information is available. A preferable alternative is where covariate information is available, for which universal block kriging or kriging with external drift would be suited. Universal kriging may be described as some spatial process which comprises both stochastic and deterministic components, and represented is by the general model:

$$Z(x) = \sum_{k=0}^{K} a_k f_k(x) + \varepsilon(x)$$

[5.3.17]

The deterministic component is represented in the above equation by a set of functions (usually first or second order polynomials), $f_k(x)$, k= 0, 1, …, K, and unknown coefficients a_k which need to be estimated based on the relationship between the target variable and covariates. The $\varepsilon(x)$ term is the stochastic field with zero mean. A block universal kriging estimate of a target variable centred by a point x_0 based on n point observations at neighbouring sites is:

$$Z_H(x_0) = \sum_{k=0}^{K} \sum_{i=0}^{n} a_k \lambda_i f_k(x_i)$$

[5.3.18]

where λ_i are the kriging weights. More detail regarding universal kriging can be found in Webster and Oliver (2001).

These particular forms of convolution (where *G* decreases and *B* increases) may also be achieved indirectly by means of performing the fine-gridding (P1→P3) and convolution as previously described (P3→P4), independently. In this example the original point support map is firstly fine-gridded using methods previously described which may or may not involve using covariate information. The second step involves increasing the support without changing the grid spacing, which was described earlier by means of block kriging. This is the logical route to follow; the alternative route i.e. P1→P2→P4 involves in addition to a decrease in resolution, two changes of support (with one step involving disseveration), which would not only be unnecessary but also a complex procedure and difficult to validate with each independent step.

5.3.2.4. Deconvolution problems

Deconvolution is a downscaling problem which involves a decrease in support such as acquiring point estimates from areal information. Digital soil mapping examples for deriving point estimates from areal information are not uncommon in soil map disaggregation exercises where a map producer will require some method to discretise points within polygons prior to generating soil attribute maps (Goovaerts 2011). Of the range of methods available, Area-to-Point kriging (Kyriakidis 2004) would be a suitable method for the deconvolution of raster-based digital soil maps. Area-to-Point (AtoP) kriging is essentially the counterpart of block kriging in that point estimates are obtained from areal (block) measurements. In the case of digital soil map deconvolution, each pixel is a block where the pixel value is some spatially averaged estimate of the target variable. The idea of AtoP kriging for deconvolution is therefore to use this areal information to discretise point estimates on a regular grid spacing as defined by the map producer. The AtoP kriging estimate for any given point x_0 is expressed as:

$$Z_{AtoP}(x_0) = \sum_{i=1}^{n} \lambda_i \cdot Z(v_i)$$

[5.3.19]

where *n* is typically smaller than the total number of block or pixels; for example (*n-1*) is the number of blocks adjacent to the block v_b where the point estimation is required. A key property of AtoP kriging is that it preserves the mass-balance or

pycnophylactic property of the blocked data; such that the average of all discretised points within each v_b returns the areal value of $Z(v_b)$. However, the constraint of equation 5.3.19 is that the same *n* areal data are used for prediction at each location within the block v_b where the point estimations are required (Goovaerts 2011). Furthermore, because areal or block estimates are used to derive point predictions, there is a requirement to know the point support variogram model. Obviously this is not available so it must be evaluated, which is done in two steps: 1) compute and model the variogram of the areal data, and 2) deconvolute the block-support model to derive the point support variogram. Goovaerts (2008) proposed an iterative deconvolution procedure that seeks the point support model that, once regularised, is the closest to the model fit of the areal data. A more detailed explanation of AtoP kriging is given in Kyriakidis (2004).

5.3.3. Conflation and Disseveration

Conflation and disseveration are strictly procedures to deal with problems for moving up and down scales when the data has some sort of areal support i.e. block-to-block processes. Conflation and disseveration procedures deal strictly with processes where the support and the grid spacing are equal and both are changed equally and simultaneously. In accordance with Figure 5.2.2 conflation processes require scaling from P2→P4. A conflation process would be carried out where given a large project area extent a map producer requires regional predictions of a target variable using available fine scaled areal estimates such as those derived for farm scales. Conflation here is a trivial upscaling problem and involves simply the averaging of the finer scaled areal observations within each coarser scaled block. With this upscaling procedure, while the overall mean of the target variable across the same map extents will remain unchanged, the overall variance will decline as the block and grid spacing simultaneously increase. The decline in variance will with increasing resolution result in the creation of homogeneous maps (Jelinski and Wu 1996).

Disseveration procedures are more complex and equate to those processes requiring scaling from P4→P2. Complexity here is due to the requirement that in addition to needing a method for estimating the variation of the target variable at a fine resolution—given that only the value at the coarse resolution is known—there is

a need to maintain the mass balance of the pycnophylactic property. Such that the target variable value given for each coarse grid cell equals the average of all target variable values at the fine scale in each coarse grid cell. This additional requirement of mass-preservation is the explicit difference between downscaling methods that involve simply fine-gridding which is essentially a points-to-points procedure and those which involve disseveration which is a block-to-block procedure.

An example of where disseveration would be enacted would be in a situation where regional estimates of a target variable (at block support) are available only at a coarse resolution and there is a requirement to generate estimates of this property to a farm or even field scale. It is therefore quite reasonable to expect that downscaling here also involves a reduction of the areal extent in addition to reduction of the spatial resolution.

More is discussed about disseveration in Chapter 6 where a suite of possible methods is detailed. Also in Chapter 6 a novel procedure for disseveration, suitable for DSM, is introduced which uses available gridded environmental information or covariates.

5.3.4. Further solutions for scaling problems

The scaling processes described so far are specific for a given problem. For example, block kriging for P1→P4 processes (convolution), point kriging for P3→P1 processes (fine-gridding), area-to-point kriging for P4→P1 processes (deconvolution) and disseveration for P4→P2 processes. Recently however, Gotway and Young (2007) introduced a generalised geostatistical framework that in addition to solving the problems of scaling described in this chapter can also be implemented for other problems that cannot be visualised by using the contrived soil maps on Fig. 5.2.2. For example, deconvolution problems that are block-to-block processes or problems that involve overlapping supports (which are described as side-scaling problems). The idea of Gotway and Young (2007) is that data of any kind of support whether it be point or block, is $Z(\mathbf{B}) = Z(B_1),\ldots, Z(B_n))$ and prediction of $Z(A)$ is of interest. The volumes A and $\mathbf{B}$ can be general which allows for several different types of scaling problems. For example, block kriging is a special case of this method when A is a volume or area and B_i are points. If A is a point and B_i is a volume or area where A is nested within B_i, the problem becomes one of deconvolution and the principals of mass-balance are preserved. Gotway and Young (2007) detail the

statistical inference of this framework and its demonstration of use. One advantage of this framework are its versatility for solving a range of scaling problems with one method, negating the requirement for resorting to specific solutions for a given problem. Furthermore, measures of uncertainty can be obtained for the predictions.

5.4. Validation of soil information products generated from re-scaling methods

The validity of the soil information that is generated from the discussed scaling processes is particularly important in the context that the information may be used for decision making or modelling purposes. Ideally the most accurate information that is available is desirable. There is certainly greater value and perceived accuracy if soil survey is coupled directly to the requirements of a project, yet in most situations this is prohibitive due to time and cost constraints. In the absence of such new soil survey, the methods discussed and the subsequent outputs will in most situations be the best available. Therefore, caution or at least awareness is required of the uncertainties associated with the new information outputs resulting from performing these scale manipulations.

Firstly, existing soil information will have some quantifiable uncertainty. As a result this uncertainty will also propagate through to the new information outputs. More often than not, the uncertainties are very rarely included with a map to be scaled, which is not ideal but something that must be acknowledged. When uncertainties are available, incorporating them into the scaling process is desirable. For example, while only demonstrated for block kriging (Section 5.3.2.2), the kriging equations can be modified to include uncertain data and this is applicable to all forms of other kriging. Furthermore, as described in the next chapter, the method used for disseveration also allows one to incorporate uncertainties of the source map into the process for creating the destination map.

Secondly, a general assumption of the methods discussed is that the behaviour of soil at large scales is explained by the average of the soil behaviour at the fine scales. This may or may not be upheld in reality or may only be relevant at a specific range of scales (Addiscott and Mirza 1998). Grunwald et al (2011), citing deYoung et al. (2008), does explain however that nonlinear dynamics and alternate states are well known in ecological systems, yet they have been poorly investigated in the soil science domain. It is beyond the scope of this chapter, but one way to investigate the

variance of soil properties at different scales, as Pettitt and McBratney (1993) suggests, is by performing nested sampling which will additionally help recognise the existence of natural hierarchies. Lark (2005) also described the value of nested sampling for understanding soil processes at different scales. Understanding the dynamics of soil processes better at different scales will obviously complement efforts when scaling of existing soil information is required.

Assessing the level of uncertainty of the new soil information outputs will ultimately be required with some form of validation. When kriging operations are performed, the kriging prediction error provides a quantitative and spatially explicit measure of the uncertainty. Otherwise, internal validations from diagnostic measures such as the coefficient of determination or the root mean square error among others provide some way of assessing the validity of outputs. This would be the case if a *scorpan* model was used for a fine-gridding operation; similarly for disseveration with covariate information. However these internal validations may be susceptible to bias (Brus et al. 2011) and the kriging prediction variances as exemplified in the example of convolution (P1→P2) will underestimate or oversimplify the true prediction uncertainty.

An unbiased approach to validating the new maps would be to collect additional samples from the study area under evaluation. To validate digital soil maps, Brus et al. (2011) recommend a design-based sampling strategy involving probability sampling. While an extra time and cost impost, design-based (and unbiased) estimation of map quality measures can be obtained with this type of sampling. When the support of observations is a point the external validation is not a technically difficult exercise. However, a technical and almost certainly costly challenge is the requirement to validate maps which have some areal support. This would require, if probability sampling is used, sample collection at a limited number of point locations within randomly selected validation supports. The average of the soil variable at these locations would be assumed as representative of the entire support size unit. Determining the number of samples for a given support unit will probably come down to some cost criterion; however, coming up with an optimal and efficient scheme for validating block support maps is an area of work will require further research.

5.5. Concluding remarks

Upscaling and downscaling are loosely used terms in the DSM community to refer to procedures for manipulating the scale of soil information outputs. However, these terms may be too general for digital soil maps because manipulations of scale not only involve changes to extent and resolution but more often than not, adjustment of the prediction support as well. I used a generic soil map model as a means to introduce terminology that explicitly describes specific scale manipulations involved in terms of changes to grid cell spacing and prediction support.

Fine-gridding and coarse-gridding are operations where the grid spacing changes but support remains unchanged. Deconvolution and convolution are situations where the support always changes which may or may not involve changing the grid spacing. Disseveration and conflation operations occur when the support and grid size are equal and both are then changed equally and simultaneously. A non-exhaustive suite of pedometric methods are described with examples of how to implement each scaling procedure. The reliability of outputs generated from scale manipulations will contain some quantifiable measure of uncertainty attributed to incomplete knowledge of soil processes at different scales and general modelling based uncertainties. It is recommended that if resources are available, validation of new map outputs is performed by collecting additional samples. Some technical challenges, particularly for validation of block support maps, persist and need to be investigated further.

5.6. References

Addiscott, T.M., 1993. Simulation modelling and soil behaviour. Geoderma, 60: 15-40.

Addiscott, T.M., Mirza, N.A., 1998. New paradigms for modelling mass transfers in soils. Soil & Tillage Research 47, 105-109.

Bierkens, M.F.P., Finke, P.A. and de Willigen, P., 2000. Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bishop, T.F.A., McBratney, A.B. and Whelan, B.M., 2001. Measuring the quality of digital soil maps using information criteria. Geoderma, 103: 95-111.

Bloschl, G., 2005. Statistical upscaling and downscaling in hydrology. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Wiley and Sons, Chichester, West Sussex, England.

Box, G.E.P. and Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society Series B-Statistical Methodology, 26: 211-252.

Brus, D.J. and de Gruijter, J.J., 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma, 80: 1-44.

Brus, D.J., Kempen, B. and Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. European Journal of Soil Science, 62: 394-407.

Burgess, T.M. and Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties. 2. Block kriging. Journal of Soil Science, 31: 333-341.

Christensen, W.F., 2011. Filtered Kriging for Spatial Data with Heterogeneous Measurement Error Variances. Biometrics, 67: 947-957.

Cressie, N., 1996. Change of support and the modifiable areal unit problem. Geographical Systems, 3: 159-180.

Cressie, N. and Wikle, C.K., 2011. Statistics for Spatio-Temporal Data. John Wiley and Sons, Hoboken, New Jersey.

Delhomme, J.P., 1978. Kriging in the hydosciences. Advances in Water Resources, 1: 251-266.

deYoung, B., Barange, M., Beaugrand, G., Harris, R., Perry, R.I., Scheffer, M. and Werner, F., 2008. Regime shifts in marine ecosystems: detection, prediction and management. Trends in Ecology & Evolution, 23: 402-409.

Finke, P.A., Bouma, J. and Hoosbeek, M.R.E., 1998. Soil and Water Quality at DIfferent Scales. Kluwer, Dordrecht.

Fuentes, M. and Raftery, A.E., 2005. Model evaluation and spatial interpolation by Bayesian combinations of observations with outputs from numerical models. Biometrics, 61: 36-45.

Goovaerts, P., 2008. Kriging and semivariogram deconvolution in the presence of irregular geographical units. Mathematical Geosciences, 40: 101-128.

Goovaerts, P., 2011. A coherent geostatistical approach for combining choropleth map and field data in the spatial interpolation of soil properties. European Journal of Soil Science, 62: 371-380.

Gotway, C.A. and Young, L.J., 2002. Combining incompatible spatial data. Journal of the American Statistical Association, 97: 632-648.

Gotway, C.A. and Young, L.J., 2007. A geostatistical approach to linking geographically aggregated data from different sources. Journal of Computational and Graphical Statistics, 16: 115-135.

Grunwald, S., Thompson, J.A. and Boettinger, J.L., 2011. Digital Soil Mapping and Modelling at Continental Scales: Finding Solutions for Global Issues. Soil Science Society of America Journal, 75: 1201-1213.

Hoosbeek, M.R. and Bryant, R.B., 1992. Towards the quantitative modelling of pedogenesis- a review Geoderma, 55: 183-210.

Isaaks, E.H. and Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jelinski, D.E., Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology 11, 129-140.

Kyriakidis, P.C., 2004. A geostatistical framework for area-to-point spatial interpolation. Geographical Analysis, 36: 259-289.

Lagacherie, P., 2008. Digital soil mapping: a state of the art. In: A.E. Hartemink, A.B. McBratney and M.L. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer Science, Australia, pp. 3–14.

Lark, R.M., 2005. Exploring scale-dependent correlation of soil properties by nested sampling. European Journal of Soil Science, 56: 307-317.

McBratney, A.B., 1998. Some considerations on methods for spatially aggregating and disaggregating soil information. Nutrient Cycling in Agroecosystems, 50: 51-62.

McBratney, A.B., Mendonca-Santos, M.L. and Minasny, B., 2003. On digital soil mapping. Geoderma, 117: 3-52.

McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S. and Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma, 97: 293-327.

Miklos, M., Short, M.G., McBratney, A.B., Minasny, B., 2010. Mapping and comparing the distribution of soil carbon under cropping and grazing management practices in Narrabri, north-west New South Wales. Australian Journal of Soil Research 48: 248-257.

Nguyen, H., Cressie, N. and Braverman, A., 2010. Spatial statistical data fusion for remote-sensing application. Technical Report No. 849. Department of Statistics, The Ohio State University, Columbus.

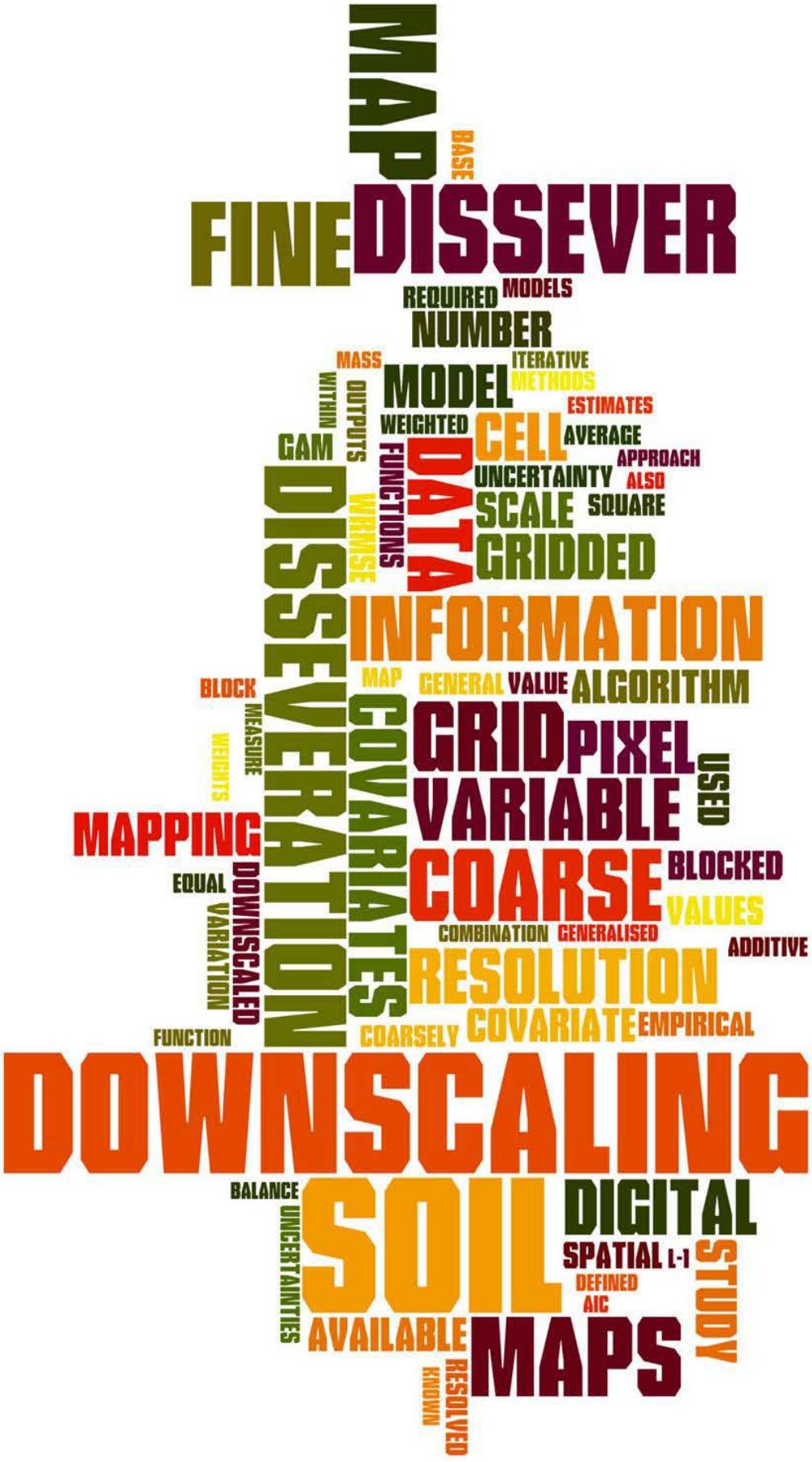
Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes - heterotopic cokriging and regression kriging. Geoderma, 67: 215-226.

Papritz, A., Herzig, C., Borer, F. and Bono, R., 2005. Modelling the spatial distribution of copper in the soils around a smelter in northeastern Switzerland. In: P. Renard, H. Demougeot-Renard and R. Froidevaux (Editors), Geostatistics for Environmental Applications Springer-Verlag, Berlin Heidelberg, pp. 343-354.

Webster, R. and Oliver, M.A., 2001. Geostatistics for Environmental Scientists. John Wiley and Sons Ltd, West Sussex, England.

Western, A.W. and Bloschl, G., 1999. On the spatial scaling of soil moisture. Journal of Hydrology, 217: 203-224.

Young, L.J. and Gotway, C.A., 2007. Linking spatial data from different sources: the effects of change of support. Stochastic Environmental Research and Risk Assessment, 21: 589-600.


MAP
BASE
FINE
DISSEVER
REQUIRED
MODELS
NUMBER
MASS
ITERATIVE
METHODS
MODEL
WITHIN
OUTPUTS
ESTIMATES
WEIGHTED
CELL
AVERAGE
GAM
FUNCTIONS
DATA
APPROACH
UNCERTAINTY
ALSO
DISSEVERATION
WRMSE
SCALE
SQUARE
GRIDDED
INFORMATION
MAP
GENERAL
VALUE
ALGORITHM
BLOCK
MEASURE
COVARIATES
GRID
PIXEL
USED
WEIGHTS
VARIABLE
MAPPING
COARSE
BLOCKED
EQUAL
DOWNSCALED
VALUES
COMBINATION
GENERALISED
ADDITIVE
VARIATION
RESOLUTION
COVARIATE
EMPIRICAL
FUNCTION
COARSELY
DOWNSCALING
BALANCE
SOIL
DIGITAL
SPATIAL
I-I
STUDY
UNCERTAINTIES
DEFINED
AIC
AVAILABLE
MAPS
RESOLVED
KNOWN

Louise: *How did you get here?*

Johnny: *Well, basically, there was this little dot, right? And the dot went bang and the bang expanded. Energy formed into matter, matter cooled, matter lived, the amoeba to fish, to fish to fowl, to fowl to frog, to frog to mammal, the mammal to monkey, to monkey to man, amo amas amat, quid pro quo, memento mori, ad infinitum, sprinkle on a little bit of grated cheese and leave under the grill till Doomsday.*

[Motion picture: Naked (1993)]

Chapter 6

A general method for downscaling digital soil maps

Summary

This chapter continues the investigations of scale manipulations for digital soil mapping, but exclusively in terms of the previously defined disseveration. In this chapter a novel methodology (called dissever*) for downscaling coarsely resolved digital soil maps using available finely gridded covariate data is presented. Under the assumption that the relationship between the target variable being downscaled and the available covariates can be non-linear,* dissever *uses weighted Generalised Additive Models (GAMs) to drive the empirical downscaling function. An iterative process of GAM fitting and adjustment attempts to optimize the downscaling to ensure that the target variable value given for each coarse grid cell equals the average of all target variable values at the fine scale in each coarse grid cell. A number of outputs needed for mapping results and diagnostic purposes are automatically generated from* dissever*. The functionality of* dissever *is demonstrated by downscaling a soil organic carbon (SOC) map with 1km by 1km grid resolution down to a 90m by 90m grid resolution using available covariate information derived from a digital elevation model, Landsat ETM+ data, and airborne gamma radiometric data.* dissever *produced high quality results as indicated by a low weighted root mean square error between averaged 90m SOC predictions within their corresponding 1km grid cell (0.82 kg* m^{-3}*). Additionally, from a concordance between the downscaled map and another map created using digital soil mapping methods there was a strong agreement (0.94). Future versioning of* dissever *will investigate quantifying the uncertainty of the downscaled outputs.*

6.1. Introduction

As described in Chapter 5, the spatial scale at which digital soil information is required is often mismatched to the scale at which it is available. One way of harmonising the 'what is required' with the 'what is available' is the application of either upscaling or downscaling methods. The focus of this study is on the application of a general method for downscaling- namely disseveration. Scale for digital soil maps as previously defined, is described in terms of extent, grid cell resolution and support. Thus downscaling in the general sense can be defined as a process involving the transfer of information from a coarser to a finer scale or resolution by either mechanistic or empirical functions (Bierkens et al. 2000).

Downscaling has particular traction in climatology research (IPCC 2001) where outputs of climate simulations from general circulation models (GCMs) cannot be directly used for hydrological impact studies of climate change because of a scale mismatch (Wilby et al. 1998; Bloschl 2005). The grid resolution of GCMs is generally in the order of hundreds of kilometres. In contrast, the resolution at which inputs to hydrological impact models are needed is in the order of tens of square kilometres. Studies by Schomburg et al. (2010) and Wilby and Wigley (1997) detail a number of approaches for downscaling GCM model output for use in driving finer scaled soil-vegetation-transfer or hydrological models. In other related environmental research fields, Liu and Pu (2008) aimed to enhance land surface temperature products using coarsely resolved satellite thermal infrared imagery. The statistical method for downscaling used by Liu and Pu (2008) was originally developed for disaggregating zonal census counts by Harvey (2002). Both Merlin et al. (2009) and Yu et al. (2008) set about downscaling soil moisture data retrieved from remote passive-microwave radiometer systems to finer resolutions in order to generate more compatible input for land surface and climate modelling. McBratney (1998) also discussed a number of potential applications for downscaling with particular reference to soil information.

Most downscaling methods can be categorized into two classes: empirical or mechanistic. Generally for either class, the problem of downscaling involves reconstructing the variation of a property at a fine resolution, given that only the value at the coarser resolution is known (Bierkens et al. 2000). Earlier studies from Tobler (1979) and, more recently, Gotway and Young (2002) detail the downscaling

approach that maintains the mass balance with the coarse scaled information known as the equal-area or pycnophylactic property. These could be simplified as approaches which attempt to harmonise the arithmetic average of the property values at the fine scale with the single property value at the coarse scale.

Linear functions, splines and general additive models are examples of empirical methods and can be exemplified by Ponce-Hernandez et al. (1986) who developed a one-dimensional mass-preserving spline method for disaggregating soil horizon data to give a continuous function of the target variable with depth. Mechanistic approaches have had considerable application in climatology research where deterministic regional climate models are nested into GCMs, which means the initial and boundary conditions to drive the regional climate model are taken from the GCMs (Yarnal et al. 2001). A popular sub-class of empirical and mechanistic downscaling approaches involves using auxiliary or covariate information (Wilby and Wigley 1997). An implicit assumption when using this auxiliary information is that they are strongly related to the target variable which is being derived at the fine scaled resolution (for examples see Schomburg et al. 2010; Bierkens et al. 2000; Wilby and Wigley 1997).

The general method presented in this study uses available fine-gridded covariate data to drive the downscaling procedure. Essentially this procedure is empirical and, through iterative model fitting, attempts to maintain mass balance; but rather than assuming a linear relationship between the target variable and the covariate data, it is also possible that the relationship can be non-linear. Therefore, a generalised multiple regression approach which replaces linear combinations of the predictors or covariates with combinations of nonparametric smoothing or fitting functions is used. This can be achieved with the use of generalised additive models (Hastie and Tibshirani 1990). Secondly, it is assumed that there is an element of uncertainty in the target variable that is being downscaled. Currently, while downscaling as a procedure is well-established (Wilby and Wigley 1997), it is often assumed or implied that there is no associated uncertainty in the values that are being downscaled. This is not the case with digital soil maps, for which there are a number of sources of uncertainty that are propagated through to the outputs. Alternatively, a soil map could be the product of a measurement or sensing device for which their will be some quantifiable measurement and or instrument error. To handle these uncertainties in the empirical downscaling process, higher weighting is given to

information which is more accurate than to information that is less accurate. I present this downscaling method as a program and subsequent algorithm called *dissever* and demonstrate its use in the downscaling of a coarse soil organic carbon map (SOC) to a finely gridded resolution.

6.2. Material and Methods

6.2.1. Algorithm for downscaling

A two-stage algorithm, initialisation and iteration, is used to downscale existing coarsely resolved target variable data to a finer resolution and support size, which is determined by the resolution of the available fine gridded environmental covariates. The algorithm presented in this study is based on that described in Liu and Pu (2008); but has been modified to accommodate the inclusion of target variable uncertainties in addition to functionality for modelling non-linear relationships between a target variable and available covariates. The algorithm is called *dissever*; meaning disseveration. Disseveration as described in Chapter 5 is a downscaling procedure where the support and the grid spacing are equal and both are changed equally and simultaneously.

The target variable value at each coarse resolution grid cell is defined as $\hat{T}_k$, k = 1, …,B; thus B is the total number of coarsely resolved grids cells across the extent of a particular study area. While $\hat{t}_m$, m = 1, …,D denotes the estimate of the target variable at each grid cell at the fine scale. In the spatial context there would be many m encapsulated by each k, the number of which would be determined by the resolution of m and will not be consistently equal, for example in study areas with non-symmetric boundaries. The number of m encapsulated by each k is denoted as E. For the initialisation stage where the iteration counter l is set to 0, $\hat{t}_m^l$ is set equal to the value of its encapsulating target variable $\hat{T}_k$. A weighted non-linear regression model between $\hat{t}_m^l$ and the suite of available covariates is fitted to all the grid cells. *dissever* uses a weighted generalised additive model (Hastie and Tibshirani 1990):

$$\hat{t}_m = \alpha + f_1(x_1) + f_2(x_2) + \cdots + f_p(x_p) \qquad \textbf{[6.2.1]}$$

where α is a constant, $x_1, x_2, \ldots, x_p$ are each of the covariate data sources, while f_j are non-parametric smoothing splines that relate $\hat{t}_m$ to the covariates. The model assumes that $\hat{t}_m$ is an additive combination of nonlinear functions of the covariates. Equation 1 can be re-written in the form:

$$\hat{t}_m = \alpha + \sum_{j=1}^{p} f_j\left(x_j\right)$$

[6.2.2]

Through an iterative back-fitting algorithm all f_j are computed which are obtained by means of a smooth of the dependent variable $\hat{t}_m$ against the covariates x_j. Justification of the back-fitting algorithm is given by the penalised residual sum of squares (PRSS) criterion which through subsequent iterations of the back-fitting algorithm is minimised (Hastie et al., 2001). Essentially the PRSS can be considered as a smoothing spline approach to estimate the additive model and is defined as:

$$PRSS\left(\alpha, f_1, f_2, \ldots f_p\right) = \sum_{m=1}^{D} w_k \cdot \left\{\hat{t}_m - \alpha - \sum_{j=1}^{p} f_j\left(x_{mj}\right)\right\}^2 + \sum_{j=1}^{p} \lambda_j \int \left\{f_j''(t)\right\}^2 dt_j$$

[6.2.3]

Each of the functions f_j is a cubic spline in the covariate x_j, with knots at each of the unique values of x_{mj}, $m = 1, \ldots, D$. The first term measures the "goodness of data fitting" or fidelity, the second term punctuated by the lambda λ_j term means 'penalties' and is defined by the functions' curvatures $\int \left\{f_j''(t)\right\}^2 dt_j$. The λ_j is considered the tuning parameter which controls the trade-off between the fidelity term and the penalties. Lastly, w_k is the weighting vector assigned to each $\hat{t}_m$. See Hastie et al. (2001) for further elaboration of the PRSS. The weighting vector of the GAM is a measure of the uncertainty that exists or was estimated in the predictions at the coarse resolutions. On the presumption that if the uncertainty is known, the weights for each grid cell k (w_k) are simply a vector where the highest weighting is given to $\hat{t}_m$ values that are the most accurate and so forth. In this study, the weights are the reciprocals of the variances of the coarse-scale grid-cell means.

dissever then shifts to the iteration stage. At the *l*-th iteration, in order to make the average of $\hat{t}_m^l$ estimates of finer resolution grid cells equal to the value of their

encapsulating coarse resolution grid cell (i.e. to equal $\hat{T}_k$), $\hat{t}_m^{l-1}$ are updated to $\hat{t}_m^l$ using the equation:

$$\hat{t}_m^l = \hat{t}_m^{l-1} \times \frac{\hat{T}_k}{\frac{1}{m}\sum \hat{t}_m^{l-1}}$$

[6.2.4]

For simplicity the average of $\hat{t}_m^l$ estimates ($\frac{1}{m}\sum \hat{t}_m^l$) will be denoted as $\bar{t}_k^l$. With the newly adjusted value, a new weighted non-linear regression model (GAM) between $\hat{t}_m^l$ and the suite of available covariates is fitted to all the grid cells. Iterations proceed until $\frac{1}{D}\sum|\hat{t}_m^l - \hat{t}_m^{l-1}|$ become equal to or decreases below a given stopping criterion value, *SCV* (the weights remain constant throughout). In the present study the *SCV* was set to 0.001. The algorithm *dissever* is summarised below in Figure 6.2.1.

Initialisation

1. $l = 0$;
2. Within each k, $\hat{t}_m^l = \hat{T}_k$ ($m = 1, \ldots, E$ and $k = 1, \ldots, B$);
3. Using a generalised additive model, regress $\hat{t}_m^l$ on $x_1, x_2, \ldots, x_p$ covariates with the weights w_k.

Iteration

4. $l = l+1$
5. Update the model estimates: $\hat{t}_m^l = \hat{t}_m^{l-1} \times \frac{\hat{T}_k}{\bar{t}_k^{l-1}}$
6. Using a generalised additive model, regress $\hat{t}_m^l$ on $x_1, x_2, \ldots, x_p$ covariates with the weights w_k
7. If $\frac{1}{D}\sum|\hat{t}_m^l - \hat{t}_m^{l-1}|$ is $\geq SCV$, repeat 4-6; otherwise, iteration terminated.

Figure 6.2.1. The downscaling algorithm written into the *dissever* program.

6.2.2. Downscaling using dissever

For this study, *dissever* was scripted in the **R** programming language (Ihaka and Gentleman 1996). It calls up the **R** package: **gam (**Generalised Additive Models) (Hastie 2011) for the regression steps of *dissever*. Operationally, *dissever* is structured as a function which requires two information inputs or objects: a data table

containing the target variable information, associated weights (if known) and covariate data source information; and the GAM formula (which is of a "formula" class **R** object) used for both the initialisation and iteration steps.

The form of the table is a data frame of *U* number of columns by *V* number of rows. Each row is a grid cell location within the area of study. Together all rows correspond to all the regular grid cell positions in the area of interest at the fine gridded scale. Generally, Columns 1 and 2 of the data table will correspond to the spatial coordinates. Column 3 is an ordinal data type column where each number corresponds to *k* (1, 2, 3,…*B*) from the coarsely gridded data. There is an obvious row number mismatch in order to arrange the coarse grid *k* to fit the corresponding number of rows at the fine gridded resolution. To overcome this, the coarse gridded information is fine gridded using a nearest neighbour re-sampling approach. Conceptually, this is just a matter of assigning the coarsely gridded cell values, here *k*, to each finely gridded cell it directly encapsulates in the spatial context. This fine gridding process is repeated also for the values of the target variable $\hat{T}_k$ and their weightings. The fine gridded attribute values and weightings are situated in columns 4 and 5 respectively. The remaining columns (6 to *U*) correspond to each of the covariate data sources that have been compiled for a study area.

It is up to the user to determine which combination of covariates to include in the model. The combination of which can be controlled by selecting the column names which correspond to the covariate data source required for inclusion. Once the two objects required for *dissever* are initialised, it is activated and will run until the stopping criterion is met or 100 iterations have run, which ever comes first. Once the function terminates, a number of outputs are created and used for mapping outputs and diagnostic analyses of the downscaling performance. A table containing the $\hat{t}_m$ predictions with appended spatial coordinates is created as are the estimates of the average of all fine gridded values ($\bar{t}_k^l$) within their corresponding coarse grid cell *k*. In terms of quantifying the mass balance deviation, iterative estimates of the weighted root mean square error (wRMSE) are given between $\bar{t}_k^l$ and $\hat{T}_k$, which is evaluated as the square root of the estimated weighted mean square error $\left(\widehat{wMSE}\right)$:

$$\widehat{wMSE} = \frac{1}{\sum_{k=1}^{B} w_k} \sum_{k=1}^{B} w_k \left(\hat{T}_k - \vec{t}_k^{\,l}\right)^2$$

[6.2.5]

Furthermore, there are iterative outputs from *dissever* which are essentially diagnostic measures of each GAM fit. The measures are given in terms of deviance which is similar to a residual sum of squares, and the proportion of deviance explained by each iterative GAM (1-[residual deviance / null deviance]), which is comparable to the coefficient of determination (R^2) from ordinary least squares regression. Akaike's Information Criterion (AIC) (Akaike 1973) is also generated from each GAM, and is a useful measure for comparing models of differing complexity, which for *dissever* would be adjusted (complexity) on the basis of the number and combination of covariates used for downscaling. The AIC is simply a measure of the relative goodness of fit of a model and is used for comparative purposes whereby the 'best' model is the one in which the AIC is minimised.

6.2.3. Case study

I demonstrate the use of *dissever* for downscaling a soil organic carbon (SOC) map featuring the variation of SOC (kg m^{-3}) in the top 30cm of the soil profile around Edgeroi, a 1500km^2 agricultural district in north-western NSW, Australia (30.32°S 149.78°E). This SOC map has a block support, consisting of 1 km by 1 km blocks centred on a square grid with a spacing of 1 km, hereafter referred to as the 1km blocked map. This map was downscaled to 90m by 90 m blocks centred on a square grid with a spacing of 90 m (90m blocked map).

The 1km blocked map was created for the specific purpose of demonstrating the application of *dissever*, such that it is the resultant product of a simple block averaging procedure (within 1km blocks) of an existing block support map; of the same support and grid cell spacing as the 90m blocked map, hereafter referred to as the 90m base map. The reason for this process was to build in a generalised validation whereby the 90m blocked map (that resulted from using *dissever*) could ultimately be compared with the 90m base map. Obviously in a true situation where downscaling would be necessary, such a comparison would not be possible.

Model-based methods in a digital soil mapping environment using legacy soil information and spatial interpolation procedures (McBratney et al. 2003) were used

to create the 90m base map—specifically, a regression kriging model, not too dissimilar to that used in Chapter 2 was used.. Block averaging predictions of the 90m base map into 1km blocks effectively created a product that might be obtained from a remote-sensing device, or might have been interpolated to this resolution because of a lack of predictive covariates at finer resolutions. For the downscaling, the covariates used by *dissever* were the same as those used to create the 90m base map. These included those derived from a digital elevation model (DEM): elevation, slope (degrees), mid-slope position, terrain wetness index (TWI) and incoming solar radiation; those derived from Landsat ETM+ imagery (2009) which included normalised difference vegetation index (NDVI) in addition to a series of band ratio derivatives:- band 5/band 7, band 3/band 7 and band 3/band 2; and those derived from airborne gamma-spectrometry information which included the channels that correspond to the abundances of both radiometric potassium and thorium. All covariate data sources were resolved to 90m grid cell resolution. for this study, w_k was the inversed variance of each 1km block averaged $\hat{T}_k$.

6.3. Results

The 90m base SOC map is shown on the top panel of Figure 6.3.1. Upscaling this map using the block averaging procedure resulted in the map on the second panel of Figure 6.3.1 (1km blocked map) and the block average standard errors (last panel of Figure 6.3.1).

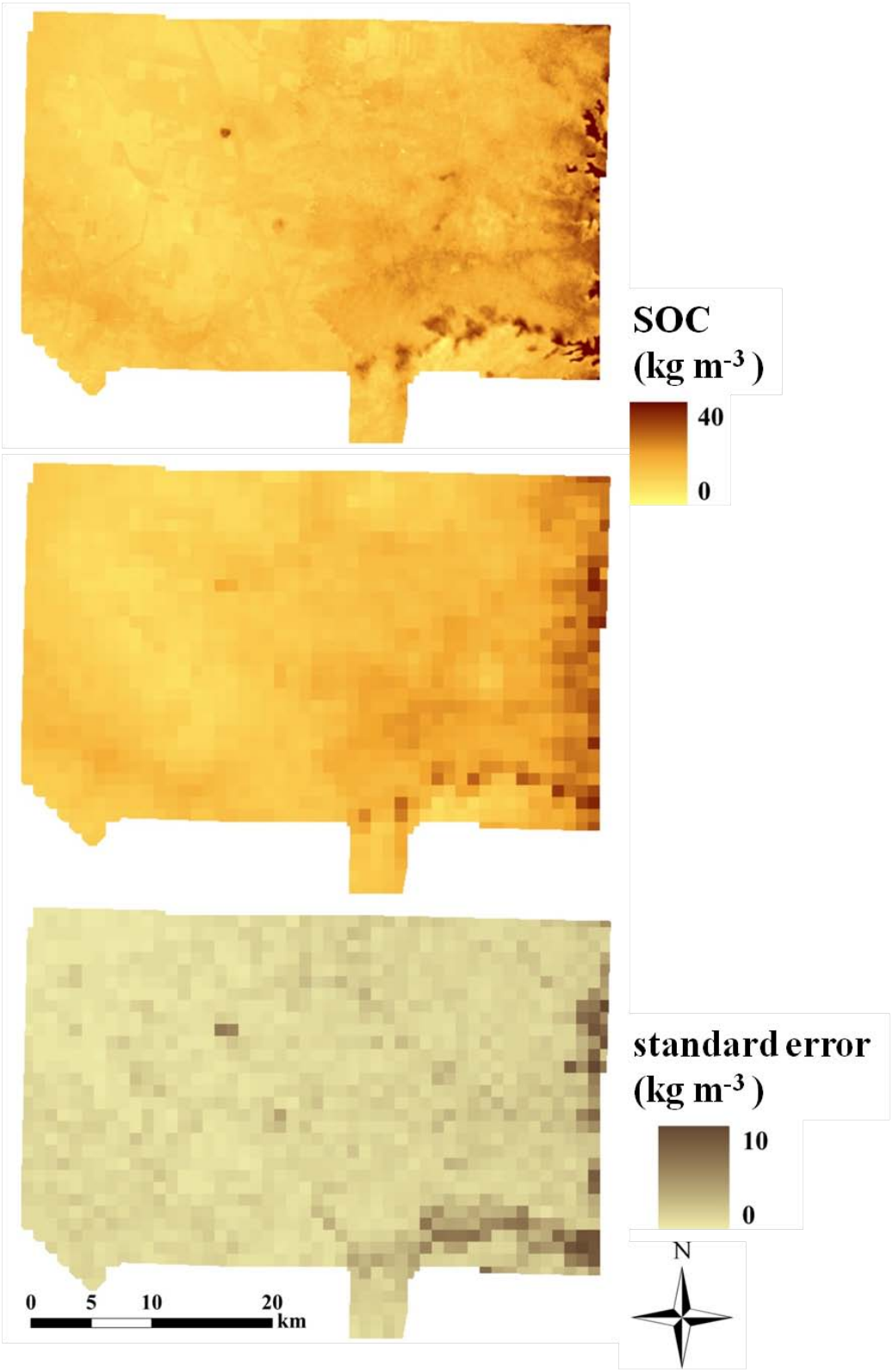


Figure 6.3.1. Top Panel: SOC map displaying the variation of SOC in the top 0-30cm across the Edgeroi study area produced from the regression kriging procedure using observed soil data and a suite of environmental covariates. Middle Panel: Upscaled map of the same target variable with 1km by 1km blocks centred onto a 1km grid produced by block averaging. Bottom Panel: Map of the standard errors of predictions resulting from the block averaging procedure.

The 90m blocked map which resulted from running *dissever* is displayed below on the top panel of Figure 6.3.2. The map on the second panel of Figure 6.3.2 is that of the absolute difference between the values of the 90m base map and the 90m blocked map, represented as two classes of difference: < 2 kg m^{-3}, and ≥ 2 kg m^{-3}. Based on these two classes about 86% (≈140 000) of the grid cells have an absolute difference of < 2 kg m^{-3}. Absolute differences ranged effectively from 0 kg m^{-3} to 8 kg m^{-3}. The third panel of Figure 6.3.2 is a plot of the comparison between both fine scaled maps. Based on this comparison there was a co-efficient of determination (R^2) of 90% (concordance: 0.94) between the soil map predictions and those resulting from the downscaling.

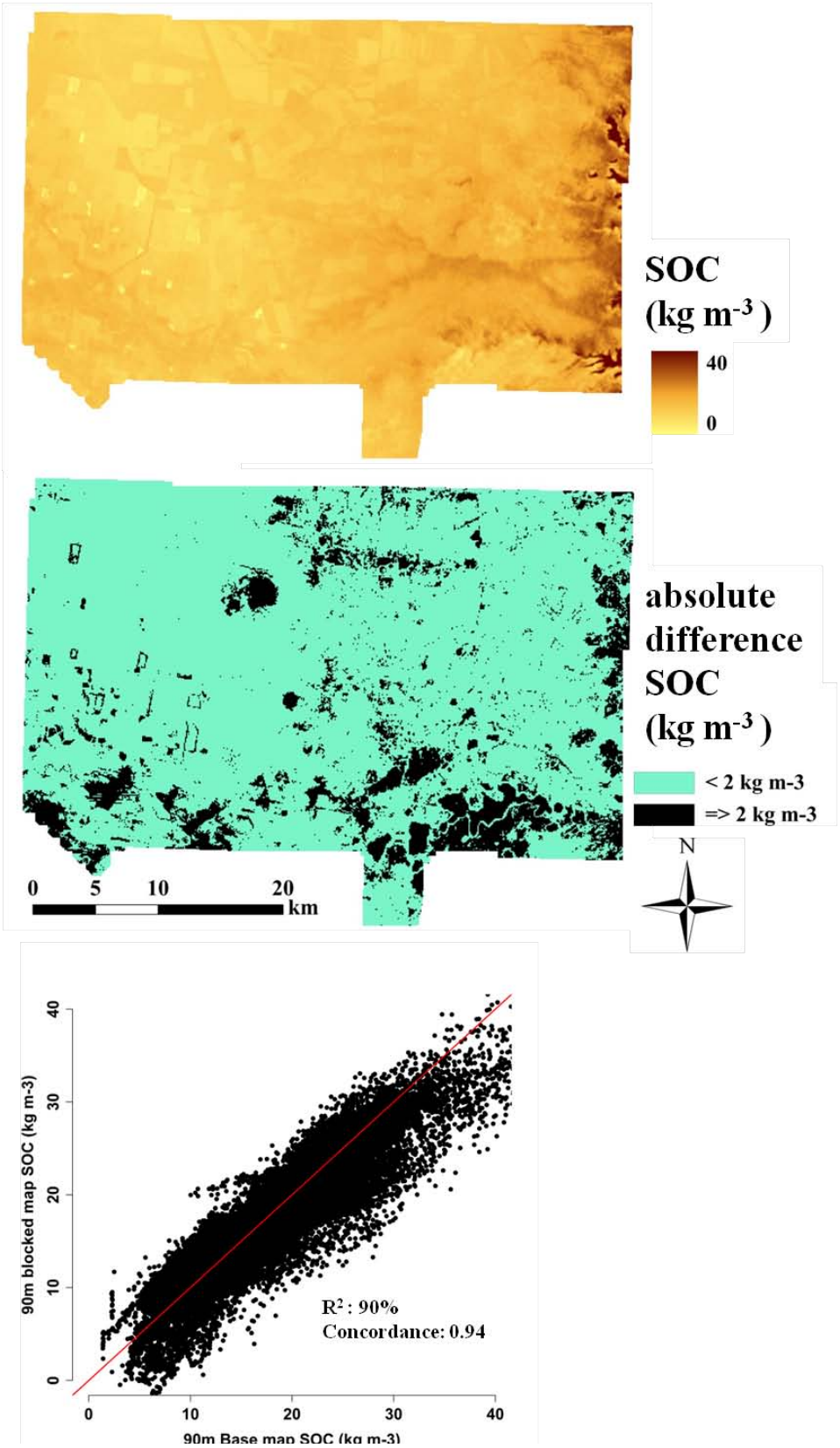


Figure 6.3.2. Top Panel: Downscaled SOC map created from *dissever*. Middle Panel: Map of the absolute differences (given as two classes of difference) between the downscaled map (90m blocked map) and the 90m base map. Bottom Panel: Concordance plot between the 90m blocked map and the 90m base map.

The real goal of downscaling is to reconstruct the variation of the target variable at a fine resolution within each coarsely resolved grid cell. To assess the quality of mass-preservation, one of the diagnostic outputs provided by *dissever* is a weighted root mean square error (wRMSE). For this case study the average deviation between the average of all fine gridded values ($\bar{t}_k^{l-1}$) within their corresponding coarse grid cell ($\hat{T}_k$) was 0.82 kg m^{-3}. In a scenario running *dissever* without incorporating the weightings on the 1km blocked map value, it was found that the wRMSE was larger at 1.10 kg m^{-3}. It was also found when running this scenario that there was a slight improvement in the R^2 (92%) and concordance (0.96) values when comparing the 90m base map with the 90m blocked map.

6.4. Discussion

The program *dissever* was designed initially for downscaling applications for DSM. However, it is a general downscaling algorithm which would suit a range of applications where scaling of information is required. This algorithm aims to determine the unknown spatial variation of a target variable at a fine resolution from an existing coarsely resolved map using a suite of finely resolved covariate or auxiliary data as predictor variables. Rather than assuming a linear function to describe the relationship between the target variable and available covariates, *dissever* makes the prediction of the target variable based on an additive combination of nonlinear functions of the covariates, which is a more general model for estimation of the unknown spatial variation. However, the GAM is not exclusive to *dissever*, and the algorithm can be simply modified to accommodate a user defined function. For example, it is possible to replace this model (GAM) with other deterministic functions which could include linear models, neural networks or regression trees as a few possibilities. While the current version of *dissever* allows the user to input the level of uncertainty associated with the information being downscaled, accommodating these uncertainties using other deterministic functions has not been investigated. In the case of *dissever* however, if the uncertainties are not known, downscaling will proceed using equal weights.

The wRMSE provides a quantitative measure to assess the mass balance deviation between the coarse gridded information and the downscaled fine gridded information, and the aim in any project is to minimise it. As discovered in this study,

taking into account the uncertainties of the 1km blocked map resulted in differing estimates of the wRMSE; 0.82 kg m^{-3} as opposed to 1.10 kg m^{-3}. In the situation of using equal weightings, the wRMSE is essentially a measure of an unweighted RMSE. The logic of including the uncertainties into the downscaling process ensures that greater weighting is given to information that is more accurate and less weighting to less accurate information; the wRMSE measure also takes this into account.

It is important to note that the wRMSE does not quantify the quality of downscaling; merely the deviation of mass balance. Thus downscaling may lead to poor results in situations where the fine grid cell variation has not been correctly predicted, even if the wRMSE is small. Nevertheless, in this study, the wRMSE appears to be quite acceptable in consideration of the concordance between the 90m base map and the 90m blocked map. This result was to be expected given that the combination of covariates used to create both maps were the same. This meets one of the implicit assumptions of downscaling using covariate data, in that they need to be strongly related to the target property which is being derived at the fine scale. The general features of both maps are comparable and where there was discrepancy it was predominantly in the order of < 2 kg m^{-3} (absolute difference).

Determining some reasons why disseveration of the coarsely resolved soil data was better in some areas than in others warrants further investigation, but is likely to have been attributed to the fact that in areas where disseveration was poorest, the uncertainty of the 1km block map was greatest. Additional to this factor, expert knowledge of the study area indicated that in areas where disseveration was poorest, there was a greater spatial variation of the 90m gridded covariate data inside each 1km block. This feature highlights a common limitation of downscaling in that irrespective of the approach, all the known variability of a target variable is seldom captured at a given scale (Wilby and Wigley 1997). It will be useful however in further research and subsequent versioning of *dissever* to determine a more sophisticated approach of assessing the uncertainties resulting from the downscaling, or in other words quantifying the confidence of the downscaled predictions. It is perceived currently that a disadvantage of *dissever* (and other downscaling procedures) is that it introduces bias attributed to differences between the averages of the fine gridded target variable data with that of the corresponding coarse gridded data. Therefore, in addition to an incomplete knowledge about the variation of the

target variable within each coarse grid cell, there is also this bias to account for when considering the magnitude of the prediction uncertainty resulting from downscaling.

With respect to the case study, the information relating to the covariates was known *a priori* to the downscaling. Such information for the general application of downscaling will obviously not be available and thus it is up to expert opinion or empirical analysis to determine suitable covariates to include in *dissever*. Empirical analysis is exemplified by the wRMSE measure in addition to the deviances and AIC estimates that result directly from the GAM fits. Particularly the AIC and to a lesser extent the residual deviance, both provide an objective tool to the user to decide which combination of covariates achieves the optimal downscaling outcome. As explained by Webster and McBratney (1989) the AIC is the statistical analogue to Occam's razor; minimising the AIC results in a fair compromise between goodness of fit and parsimony.

Overall, this program was tested on a dataset with 175 000 grid cell nodes (fine-grid). With this size dataset, downscaling terminated after 1-2 hours. However, the computational time required is dependent upon the complexity of the GAM used (increasing or decreasing the number of predictive covariates). Generally, its usefulness for downscaling has been demonstrated. There is some expertise required to arrange the spatial data to generate the input table required by this program. More importantly, however, is the necessary technical and theoretical expertise to decide which auxiliary data sources dominate at the scale for which the target variable is being downscaled.

6.5. Conclusions

One issue of spatial information is that the scale at which it is available is often inadequate or does not correspond to the scale at which it is required. There are established methods for upscaling and downscaling which are able to address these issues. The program *dissever* described in this chapter is a new program that builds on existing empirical methods of downscaling earth resource information, yet is suited specifically for DSM. Principally, while attempting to maintain the mass balance with the available coarse scaled information, *dissever*, through an iterative algorithm, attempts to reconstruct the variation of a property at a prescribed fine

resolution through an empirical function using auxiliary information. The features which differentiate it from other methods are:

- It generalises the multiple regression approach which replaces linear combinations of the predictors or covariates with combinations of non-parametric smoothing or fitting functions. This generalised fitting allows the possibility to accommodate non-linear relationships between the target variable and the covariates.
- The target variable uncertainties at the coarse scale are incorporated into the downscaling algorithm which subsequently moderate the outcomes of the downscaled products and associated measures of mass balance deviation.

6.6. References

Akaike, H., 1973. Information theory and an extension of maximum likelihood principle. In: B.N. Petrov and F. Csaki (Editors), Second International Symposium on Information Theory. Akademia Kiado, Budapest, pp. 267-281.

Bierkens, M.F.P., Finke, P.A. and de Willigen, P., 2000. Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bloschl, G., 2005. Statistical upscaling and downscaling in hydrology. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Wiley and Sons, Chichester, West Sussex, England.

Gotway, C.A. and Young, L.J., 2002. Combining incompatible spatial data. Journal of the American Statistical Association, 97: 632-648.

Harvey, J.T., 2002. Population estimation models based on individual TM pixels. Photogrammetric Engineering and Remote Sensing, 68: 1181-1192.

Hastie, T.J., 2011. Generalised Additive Models: R Package 'gam'. R Foundation for Statistical Computing, Vienna, Austria.

Hastie, T.J. and Tibshirani, R.J., 1990. Generalized Additive Models. Chapman and Hall, London, England.

Hastie, T.J., Tibshirani, R.J. and Friedman, J., 2001. The Elements of Statistical Learning: Data mining, inference and prediction. Springer, New York, NY.

Ihaka, R. and Gentleman, R., 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics, 5: 299-314.

IPCC, 2001. Climate Change 2001: The Scientific Basis. Contribution of the Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York.

Liu, D.S. and Pu, R.L., 2008. Downscaling thermal infrared radiance for subpixel land surface temperature retrieval. Sensors, 8: 2695-2706.

McBratney, A.B., 1998. Some considerations on methods for spatially aggregating and disaggregating soil information. Nutrient Cycling in Agroecosystems, 50: 51-62.

McBratney, A.B., Mendonca-Santos, M.L. and Minasny, B., 2003. On digital soil mapping. Geoderma, 117: 3-52.

Merlin, O., Al Bitar, A., Walker, J.P. and Kerr, Y., 2009. A sequential model for disaggregating near-surface soil moisture observations using multi-resolution thermal sensors. Remote Sensing of Environment, 113: 2275-2284.

Ponce-Hernandez, R., Marriott, F.H.C. and Beckett, P.H.T., 1986. An improved method for reconstructing a soil-profile from analysis of a small number of samples. Journal of Soil Science, 37: 455-467.

Schomburg, A., Venema, V., Lindau, R., Ament, F. and Simmer, C., 2010. A downscaling scheme for atmospheric variables to drive soil-vegetation-atmosphere transfer models. Tellus Series B-Chemical and Physical Meteorology, 62: 242-258.

Tobler, W.R., 1979. Smooth pycnophylactic interpolation for geographical regions. Journal of the American Statistical Association, 74: 519-530.

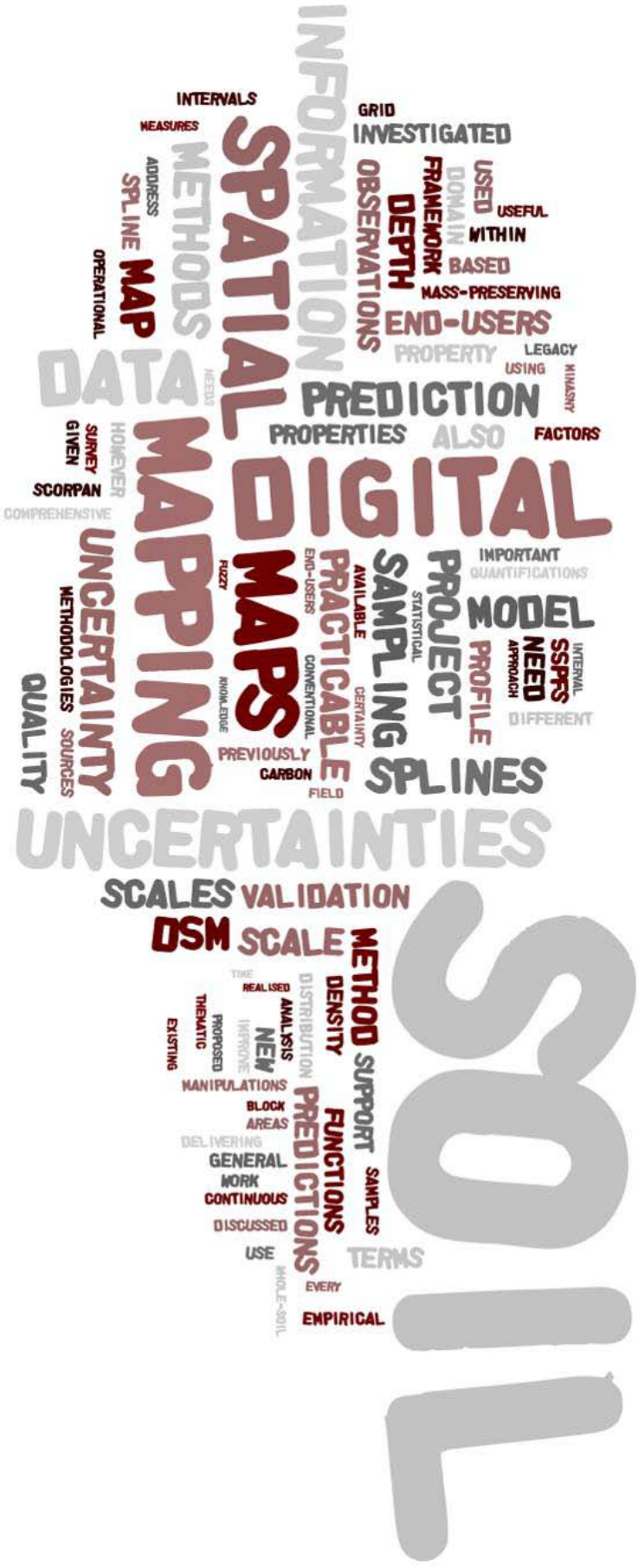
Webster, R. and McBratney, A.B., 1989. On the Akaike Information Criterion for choosing models for variograms of soil properties. Journal of Soil Science, 40: 493-496.

Wilby, R.L. and Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography, 21: 530-548.

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and Wilks, D.S., 1998. Statistical downscaling of general circulation model output: A comparison of methods. Water Resources Research, 34: 2995-3008.

Yarnal, B., Comrie, A.C., Frakes, B. and Brown, D.P., 2001. Developments and prospects in synoptic climatology. International Journal of Climatology, 21: 1923-1950.

Yu, G., Di, L. and Yang, W., 2008. Downscaling of global soil moisture using auxiliary data, 2008 IEEE International Geoscience and Remote Sensing Symposium. Institute of Electrical and Electronics Engineers, Inc., Boston, Massachusetts, U.S.A., pp. 230-232.


SOIL
INFORMATION
SPATIAL
DIGITAL
MAPPING
DATA
MAPS
UNCERTAINTIES
UNCERTAINTY
SAMPLING
PREDICTION
PREDICTIONS
METHOD
METHODS
MODEL
PROJECT
PRACTICABLE
SPLINES
SCALES
VALIDATION
DSM
SCALE
OBSERVATIONS
DEPTH
END-USERS
PROPERTIES
PROPERTY
FACTORS
ALSO
IMPORTANT
NEED
PROFILE
QUALITY
METHODOLOGIES
SOURCES
FUNCTIONS
SUPPORT
TERMS
EMPIRICAL
INTERVALS
GRID
INVESTIGATED
USED
USEFUL
WITHIN
BASED
MASS-PRESERVING
LEGACY
USING
SPLINE
MAP
OPERATIONAL
ADDRESS
MEASURES
SURVEY
GIVEN
SCORPAN
COMPREHENSIVE
HOWEVER
FUZZY
AVAILABLE
CONVENTIONAL
KNOWLEDGE
CERTAINTY
PREVIOUSLY
CARBON
FIELD
STATISTICAL
APPROACH
INTERVAL
DIFFERENT
QUANTIFICATIONS
REALISED
DISTRIBUTION
DENSITY
ANALYSIS
NEW
IMPROVE
PROPOSED
THEMATIC
EXISTING
MANIPULATIONS
BLOCK
AREAS
DELIVERING
GENERAL
WORK
CONTINUOUS
DISCUSSED
USE
EVERY
SAMPLES

Cast down your bucket where you are.
[Booker T. Washington 1895]

Chapter 7

General discussion, conclusions and future work

7.1. General discussion

7.1.1. Comprehensive spatial soil information

This thesis has set forth some practicable methodologies for delivering comprehensive spatial soil information. In terms of comprehensive spatial soil information, Chapter 2 addressed the need for whole-profile mapping of soil property distribution. Fundamental to this approach was the coupling of equal-area spline depth functions with digital soil mapping (DSM) procedures. Predictions alone however do not suffice for critical decision-making purposes; rather it is equally important to have some measure of certainty attached with the predictions in order for the end-users to determine the reliability or suitability of the spatial soil information products. Chapter 3 addressed this need by setting forth an empirical method of uncertainty analysis for DSM. This approach is independent of the soil spatial prediction function (SSPF). Uncertainty is expressed as a prediction interval, where for a nominal confidence level, at every pixel or grid cell, a reciprocal level of certainty can be given to the expectation that the true but unknown soil property value would lie within the interval. This uncertainty method was extended to quantify the prediction uncertainties for whole-soil profile mapping of soil properties. New map validation criteria were then proposed in Chapter 4 which provided quantitative measures of map quality based on the predictions *and* their uncertainties. These measures were the Mean Square Error of a Simulated random value (MSES), which is a prediction accuracy criterion that takes into account the prediction uncertainties. The second criterion was the Prediction Interval Coverage Probability (PICP) expressed as the Areal Proportion of Correct Predictions (APCP) given a 95% confidence level. The APCP explicitly addressed the quality of the quantifications of the uncertainties.

These chapters (2, 3, and 4) represent a complete framework for delivering comprehensive whole-soil profile digital soil maps that have quantified uncertainties at every grid cell and depth. Furthermore, the framework incorporates delivering quantified measures of map quality both in terms of the predictions and the

uncertainties. Admittedly these three chapters dealt with DSM at the regional or watershed spatial scales without consideration to whether this framework would apply to other spatial scales. This is a minor issue and not likely to pose problems due to the frameworks' flexible nature; there are some operational constraints nonetheless. One would be the requirement of available data, both observed soil data and environmental or secondary information (covariates) at the scale of interest to build and then subsequently extend the SSPFs and uncertainty estimations. A related constraint to extension of the framework to other spatial scales is observing and explicitly taking into account the scale-dependent behaviour of the soil-landscape relationships at the scales intended to be mapped.

Chapters 5 and 6 then proposed a framework and new tools where, without the requirement of new soil survey, digital soil maps could be created to the spatial scale specifications of the end-users from existing maps. The framework is based on methods of upscaling and downscaling that observe the scaling triplet for digital soil maps: extent, resolution, and support.

7.1.2. Practicable methodologies

The concept of practicability will differ from person to person and their level of expertise. Furthermore how we define *what is practicable* will most likely change temporally as statistical theory is advanced and efficiencies in computing are further realised. Yet, moving towards the realisation of operational DSM means that practicable methods need also accommodate for the data that is used for mapping, such that often it comes from legacy soil survey which is often not ideal in terms of spatial distribution within a mapping domain, spatial density and accuracy (regarding measurement and positional accuracies). Furthermore, practicable methods must also be general enough to suit a range of different situations where they can be applied for a number of different soil properties.

In Chapters 2, 3, and 4 there was a preference towards using SSPFs which were structurally based on regression kriging or *scorpan* kriging approaches (McBratney et al. 2003). Regression kriging is a general method that allows flexibility in selection of the trend model (deterministic model) between the observed data and the *scorpan* factors. These models may be linear or non-linear and even allow for incorporation of data-mining methods. We then treat the model residuals independently to determine whether there is any spatial structure that may be present

after de-trending, with the view that this additional spatial structure can be captured to improve the overall prediction outcomes. As previously described in the review, there are reservations about this approach in that it is theoretically biased (Cressie 1993); empirical studies suggest that in reality the method is robust to biasedness. Considerations aside, the method is practicable; first because there is flexibility for deciding which type of deterministic model is appropriate for a given project. Secondly, the approach easily handles large data sets which for more statistically sound methods such as REML-EBLUP (Lark et al. 2006) may be prohibitive or may only make small overall improvements in the model outcomes.

The equal-area spline depth function is a good example of a practicable method for modelling the depth variation of soil properties. Bishop et al. (1999) has previously shown it to be useful for a number of soil properties. In this project, splines were fitted to soil profile data describing variations in carbon storage and available water capacity (Chapter 2 and 3) and soil pH (Chapter 4) respectively without issues regarding the quality of the spline fit to the raw observed data (see plots on Figure 2.3.3 as an example). Admittedly, the raw data used in this project were of high quality where the frequency of observations recorded for each profile was generally quite high. However, the spline function is not a pedological model; rather it is a data model where the quality of the fit in terms of representing pedological features such as abrupt changes in soil behaviour is largely determined by the frequency of observations recorded within each profile. Wide application of the spline function has been demonstrated in this project, and it is probably more practicable than other depth functions proposed for DSM such as negative exponential depth functions (Minasny et al. 2006) and ones that incorporate explicit pedological knowledge (Kempen et al. 2011) as they are suited only to a few soil properties or only work in the areas where the pedological knowledge was established in the first place. It is not that these methods are inferior; rather the extent of their application is limited.

Common methods of uncertainty analysis involve Monte Carlo simulations (Heuvelink et al. 1998) or Bayesian approaches (Diggle et al. 1998), which for operational DSM could be computationally prohibitive to consider as being practicable for DSM at global scales, for applications related to precision agriculture, or any situation that involves the use of very large datasets. A pragmatic step in a relatively complex field was therefore taken in this project to propose a practicable

method for quantifying uncertainties in relation to DSM. As described previously, uncertainty in this project is treated as the probability distribution of the output model errors, which comprises all sources of uncertainty (model structure, model parameters and input data). This is particularly useful when we are dealing with SSPFs that include data-mining tools or neural networks (as examples) in combination with the regression-kriging approach, where it would be difficult to use other existing methods (of uncertainty analysis) to derive model parameter uncertainties. It is also useful because it does not use multiple realisations of simulations to generate the probability distributions. Because of these efficiencies, what this empirical uncertainty approach facilitated overall was the ability to quantify the whole-profile prediction uncertainties generated from DSM at every grid cell. Moreover, the prediction intervals were generated without being a computational burden. Validation of the prediction intervals particularly in Chapter 3, indicated the estimates of uncertainty were not unfounded whereby for a given confidence level, we would expect a corresponding proportion of observations to fall within their given prediction interval. In Chapter 4, it was, however, revealed that the prediction intervals are subject to bias and are not so representative at sub-soil depths. These deficiencies are discussed in more detail below. Technical issues aside, the fact that the quantifications of the uncertainties could be validated in the way they were in this project represents a significant step forward for DSM in general. Validation of soil maps unfortunately is not widely practiced, and validation of prediction uncertainties has seldom been done before for DSM. The impediments of computationally demanding uncertainty analysis procedures as previously discussed may be one of the constraints to wider testing of the quantifications of uncertainties. For operational DSM, application of the empirical approach for assessing uncertainties may negate some of these constraints.

For assessing soil map quality, collection of additional samples from the field is recommended and variants of random sampling are ideal because unbiased estimators are quantified (de Gruijter et al. 2006). Unfortunately the time, cost and effort factors may prohibit this additional fieldwork from being carried out. Other methods of validation as previously discussed in the review include random holdback and leave-one-out-cross-validation which require no new data; rather they require some data to be left out of the model calibration process which are then used to test its (model) estimation quality. These methods are not optimal as the map quality

estimators will often be biased because the legacy data used to generate the digital soil map often does not come from a probability sample. But they are practicable alternatives when independent sampling can not be performed. The view is that some form of validation is better than no validation, and for end users to objectively determine the reliability of a map, it is important that proper evaluation of digital soil maps is provided. Ultimately how this evaluation is done will be up to the responsible party for producing the map and what is practicable given the resource constraints.

One of the biggest impediments to operational DSM is the availability of data at the spatial scales at which mapping needs to be performed. There is no general rule prescribing the ideal sampling density for mapping at a given spatial scale. However, empirical evidence from a review of soil carbon mapping suggests that sampling density decreases logarithmically as resolution increases (Minasny et al. 2012), which is analogous to high sampling density at field and farm spatial scales to low sampling density at continental or global spatial scales. Empirically, sampling densities have ranged in the order of 0.002 to 1100 samples per km^2 for soil carbon mapping at global and field spatial scales respectively (Minasny et al. 2012). Because there is no general rule regarding sampling density and DSM, there are also no restrictions on generating maps at high resolutions using low sampling densities; the uncertainties of the predictions will reflect this however. Ideally, to generate a digital soil map one would want a representative sample from the mapping domain. Constraints due to time, cost and effort may make sampling the mapping domain an unviable option. It was in this context in Chapter 5 that a pragmatic step was taken by proposing a framework for going about scale manipulations on existing digital soil maps in order to generate spatial soil information products to the scale specifications of the end-users. The framework is based on practicable methods of upscaling and downscaling that observe the scaling triplet for digital soil maps: extent, resolution, and support. As such, upscaling and downscaling are probably too general for description, which therefore prompted the introduction of new terms: *Fine-gridding* and *coarse-gridding* for operations where the grid spacing changes but support remains unchanged. *Deconvolution* and *convolution* are operations where the support always changes, which may or may not involve changing the grid spacing. D*isseveration* and *conflation* operations occur when the support and grid size are equal and both are then changed equally and simultaneously. All of these methods

are desktop procedures which require no new soil observations. There is however, particularly for *fine-gridding*, *deconvolution* or *disseveration*, usually a requirement of covariate information that may explain the variation of soil at finer scales when working from coarse scaled digital soil maps. Fortunately, finding data of this kind is more and more likely as there is an increasing availability of data sources. Besides a variety of new environmental information, the spatial resolutions of these data are becoming available at finer resolutions (Minasny et al 2008). The availability of these new data sources is invaluable for downscaling and also useful for updating existing digital soil maps.

7.2. Summaries of research findings

7.2.1. Mapping continuous depth functions

The parameters of the spline used for DSM are the harmonised depth interval observations that result from fitting splines to the raw soil profile or core data. This means that the values of the parameters at each depth are *observations*- not in the sense of the recorded 'hard' observations rather the interpolated 'soft' observations resulting from the spline fit. When the spline is fitted acceptably to a soil profile there is little difference between these 'hard' and 'soft' observations. By harmonising a collection of soil cores or profiles within a mapping domain in terms of the prediction depth intervals, there is allowance to concentrate predictions at specific depth intervals. For example there may be a requirement to generate a map of carbon concentration for the top 5 cm across a mapping domain; using the splines we can get around the issue that maybe for most of the soil profiles or cores, there is no explicit record of carbon concentration at this interval. Alternatively, prediction can be made at a succession of harmonised depth intervals (that are defined by the modeller) which can later be used as spline parameters to create continuous soil profiles at every grid cell within a mapping domain (as was done in this project). Once a continuous soil profile is predicted at every grid cell, Chapter 2 and 3 illustrated the amenability of the resulting geo-database for user-defined queries, for example, determining the depth at which soil carbon concentration decreases to below 1% etc. Having the tools to characterise and query soils in this pseudo-3D way, new possibilities for the retrieval and enquiry of spatial soil information can be realised.

Currently one of the problems however for mapping continuous depth functions is that the SSPF accuracy decreases with depth, which has also been reported elsewhere in Minasny et al. (2006) and Kempen et al. (2011). This indicates that the *scorpan* factors commonly used for DSM mainly explain the soil conditions in the top 30 to 50 cm. Yet their power for explaining soil variations in the subsoil can be quite poor. The challenge therefore is to find *scorpan* factors that are powerful for sub-soil predictions.

7.2.2. Analysing uncertainties and the validation of digital soil maps

Much has been discussed already about the practicability of the empirical uncertainty method used in this project. There are definite advantages in terms of computational efficiency and implementation compared to other approaches to analysing uncertainties.

Fundamental to delivering comprehensive spatial soil information is attaching some quantification of the certainty to each prediction. This sort of information goes a long way towards making objective decisions about the suitability of digital soil maps to address a particular question. Digital soil maps are never made with complete certainty and thus we need to investigate why our SSPFs work well in some areas and poorly in others. These efforts then provide opportunities to improve maps in an objective way by targeting sampling to areas where predictions are poor.

The key assumption of the uncertainty method used in this project is that areas that are common in terms of their environmental characteristics within a mapping domain will share a similar range of uncertainty. The root of this assumption is based on the fact that the covariate space of the soil observations are clustered into fuzzy classes from which each will then have an associated distribution of model errors or residuals. Using the fuzzy k-means with extragrades algorithm (McBratney and de Gruijter 1992) is useful for defining fuzzy classes where the extragrades are considered the outliers of a dataset (which may have a distorting influence on the configuration of the main clusters) and exist spatially in regions of low density observations. This is important for soil mapping and analysing the uncertainties as areas that have low sampling density generally have higher prediction uncertainties which are represented as wider prediction intervals. The opportunity therefore to improve the maps generated in this project would be to target further sampling in the areas where the density of existing samples is low.

A requirement for future studies using the empirical uncertainty method is to perform the clustering of the covariate space independently for each soil property that is being investigated. It was not optimal in Chapter 3 to use the same *scorpan* factors for building the SSPFs for both available water capacity and carbon storage. This also meant that the fuzzy classes were the same for each soil property. For good DSM practice, it is important to determine what the important soil-landscape characteristics are for each target soil property. Not only will this improve the predictions alone, but the quantifications of the uncertainties will also improve. For DSM this will mean either expert decision-making to decide which *scorpan* factors are important or perform a statistical multivariate analysis. This will mean ultimately that the characterisation of the fuzzy classes used for defining the uncertainties are to be performed independently for each target soil property. This is extra work but the computational efficiencies of the uncertainty method mean that this targeted work is not prohibitive.

Ultimately, the objective way to evaluate the quality of digital soil maps in terms of the predictions and their uncertainties is through validation. With the quantification of prediction uncertainties, Chapter 4 proposed new criteria to test their quantifications- MSES and APCA. Coupling these criteria with conventional measures of map quality (which address the quality of the predictions only) such as the MSE, ME and imprecision, provides new tools for decision making in the presence of quantified uncertainties. In the end, regardless of the purported map quality, it is up to the end-user to determine the suitability for addressing a particular question.

Assessing the quality of the uncertainties with validation did however reveal that prediction intervals are susceptible to bias, which was an outcome in Chapter 4. This bias results from using legacy soil data for model calibrations which do not represent a statistical sample of the mapping domain. It is difficult to remedy this as we can only work with the data that is available. Awareness of bias however presents a valuable opportunity for improving digital soil maps by sampling in the areas where the uncertainties are highest. Spatially balanced sampling designs, such as generalised random-tessellation stratified sampling (GRTS; Stevens et al. 2004) may be adapted in these situations because they ensure the whole geographical space is satisfactorily sampled. A similar spatially-balanced design has been proposed by Walvoort et al. (2010) which uses compact geographical strata from *k*-means

classification. Alternatively, the use of other statistical sampling designs for new soil survey work may be implemented. For example, approaches such as stratified random sampling of the geographical space (McKenzie and Ryan 1999) or Latin Hypercube Sampling (Minasny and McBratney 2006) are recommended. It is with these methods that representative samples can be taken from the mapping domain as the whole geographical space (covariate space) defined by the *scorpan* factors can be sampled.

Chapter 4 focused on validation of a soil pH map, and the sampling design and characterisation of the strata were based on the prior predictions of pH and their uncertainties. In a way the sampling design was optimised for validating the soil pH map, but not for other thematic maps that may need to be validated at some other given time. It is not that the additional samples cannot be used for validation of other thematic maps; it is that the samples no longer represent a proper statistical sample in the strictest sense. A practicable way as discussed in Chapter 4 to concurrently evaluate a number of thematic soil maps is to perform compact geographical stratification, where the mapping domain is stratified on the basis of spatial coordinates, followed by a random sample from each stratum (Brus et al. 1999).

7.2.3. Comprehensive digital soil mapping using low-cost input data

From the two datasets used in this project—The Edgeroi Dataset and The Hunter Valley Soil Dataset—common features were apparent. Firstly, the soil spatial prediction functions were not overly powerful in predicting the total variation that was observed in each dataset. Secondly, from the associated uncertainty analyses for both datasets, considerable levels of uncertainty were quantified. From what has been established from this thesis, the likely cause of these outcomes is twofold. 1) Soil is incredibly complex and is difficult to comprehend with the current quantitative models and covariate data sources at the disposal for DSM —our models are merely just abstractions of the reality. 2) In terms of geographical coverage and sampling density, the input data used to calibrate the spatial soil prediction functions were not optimal given the scale at which the mapping was produced for. Given the knowledge that soil variability increases with increasing areal extents (Burrough 1993) , and the reliance of operational DSM using sparsely populated, legacy soil datasets, similar outcomes as observed in this study (possibly even worse) have been observed or are to be expected for other projects. As stated in Chapter 2, low and

accuracies and/or high estimates of uncertainty should not be interpreted as a failure of the method, rather as a limitation of the low-cost data that is used by construct the spatial soil prediction functions. The ability of DSM methods being able to quantify map accuracies and measures of uncertainty means one can express strong rationale for motivating further investments in input data—including both soil information and environmental covariate data sources.

The question is, what is an acceptable level of accuracy, or what threshold for uncertainty can one accept before a digital soil map can be deemed acceptable or unacceptable? This was touched on in Chapter 4, and basically when it comes down to it, it is the end-user whom decides the fitness for a particular purpose. This may be interpreted as a vague recommendation; however, there is no end in difficulty in determining what ones perception of acceptable is without firstly asking what purpose is the map to serve and at what scale is the information required. There is not much if any known literature which broaches the topic of defining soil map quality thresholds on the basis on who the end-users are, and at what scale the information is required. This area of specification that is likely to become important soon and into the future as DSM moves further into operational status.

Dependent on the nature of the digital soil map end-use, intuitively, what is deemed acceptable will invariably range significantly. For example, one could easily ascribe different levels of map quality if it to be used for informing policy, or determining landuse suitability, or estimating carbon stocks. On assessment, one would expect the criteria for accepting a map for the use of assessing carbon stocks to be a lot stricter (more accurate) than that for assessing landuse suitability etc.

Similar intuition is needed when assessing soil maps at different spatial scales. In situations where precision agriculture is performed (field and farm scales), there will be a lot of observed soil points with which to develop a reliable spatial prediction function. From this we would also assume the spatial variability of the target soil properties at these scales is less than at larger scales. Thus one would expect more accurate soil maps. Conversely at larger scales, where there is usually less data, yet more soil variability to capture; one would accept soil maps with purported lower accuracy and higher uncertainties (than for field and farm scale DSM) more readily.

In summary, with the low-cost data we often use for digital soil mapping, we have to live with, and be accepting of, a lot of uncertainty. One should not be in the practice of discarding soil maps. Rather, there is a compelling argument for policy

makers to invest in soil mapping efforts. If it is the end-user whom requires soil information to meet some quality criteria, then they must invest appropriately to achieve those ends, rather than relying on the low-cost information. The digital soil mapping toolbox is more than capable of meeting the needs of the end user given further investment on their part in regards to input data. There is however some work to do on putting numbers to what is acceptable from the perspective of what the soil information is to be used for, and due consideration of the mapping scale.

7.2.4. Scale manipulation

Soil survey to complement digital soil map creation is ideal but not always possible for every situation where spatial soil information is required. One of the assumptions of the scale manipulations investigated in Chapters 5 and 6 is that the behaviour of soil at large scales is explained by the average of the soil behaviour at the fine scales. From an implementation perspective, this makes the methods of scale manipulation more practicable. The assumption may or may not be upheld in reality or may only be relevant at a specific range of scales (Addiscott 2010). Non-linear relationships between different spatial scales have been observed in ecological systems and most probably will be observed in terms of soil variability at different spatial scales (Crawford 2010). Thus, end-users should be aware that outputs generated from scale manipulations will contain some quantifiable measure of uncertainty attributed to incomplete knowledge of soil processes at different spatial scales and general modelling based uncertainties. While some of the scale manipulation methods that are based on geostatistical concepts can provide measures of certainty of the predictions such as kriging variances, this project did not address methods of formal uncertainty analysis for digital soil maps resulting from scale manipulations. Some pointers to be aware of however include the need to know the uncertainties of the source map. These then need to be incorporated within the formulation of the scaling methods. The program for disseveration developed in this project- *dissever* is one such method that allows the incorporation of prior uncertainties. Alternatively, when geostatistical methods are used, the incorporation of measurement or prediction errors into the kriging system is also possible.

Like other digital soil maps, evaluation of the outputs from scale manipulations should ideally be quantified by validation, and preferably by independent sampling. An important consideration in these situations is to ensure

that the support at which validations are made are equivalent to the support of the map to be validated i.e. the entity for validation of block support maps need also to be block, similarly point support validations for point support maps. Some technical challenges particularly for validation of block support maps need to be investigated further. For example, how many samples need to be collected from a block so that it is representative of the block?

7.3. Overall research conclusions

Relevant spatial soil information is required for objective decision making to (1) address current problems associated with soil degradation; (2) for modelling, monitoring and measurement of particular soil services; (3) for general management of soil resources. Depending on the purpose, the likely spatial scales at which this information will be required, ranges from field size scales (resolution <20m) up to global scales (>2km). Predominantly using legacy soil data, this project set about proposing methodologies for delivering comprehensive spatial soil information to address these needs, with the idealistic pursuit that the methodologies be practicable:

- Comprehensive in the sense that delivered spatial soil information is tailored to meet the spatial scale specifications of the end user, and is of a nature that fully characterises the whole-soil profile with associated prediction uncertainties, and where possible, both the predictions *and* uncertainties have been independently validated.
- Practicable in the sense that the methods are general; can be applied to a wide range of soil properties; can handle variable qualities of data; and are effective when working with very large datasets. Ideally, practicable methodologies will also be computationally efficient.

The investigations carried out during this project realised the following:

- Equal-area spline depth functions are a useful tool for deriving the continuous variation of soil properties from soil profile and core observations. They are also suitable to use for a number of different soil properties. The coupling of the spline depth functions with digital soil mapping facilitates mapping in the lateral and vertical dimensions (whole-soil profile), the continuous

distribution of soil properties. With whole-soil profile digital soil maps, new vistas for the retrieval and enquiry of spatial soil information can be realised.

- An empirical method for quantifying all sources of uncertainty (model structure, model parameters and input data) propagated through SSPFs for digital soil mapping was devised. This method was applied to assessing the prediction uncertainties of whole-soil profile digital soil maps where uncertainty is expressed as a prediction interval of the underlying model errors. The method is computationally efficient and amenable for complex SSPFs such as regression kriging approaches.
- New criteria were devised to facilitate validation of the quantifications of uncertainties that result from digital soil mapping. These criteria would ideally be coupled with conventional measures of soil map quality to provide a proper evaluation of digital soil maps in terms of both the predictions *and* their uncertainties.
- A framework of upscaling and downscaling approaches for digital soil mapping was devised. The framework takes into account the scaling triplet of digital soil maps and recommends pedometric methodologies for scale manipulation based on the scale entities of the source and destination maps. Scale manipulations are suitable for tailoring spatial soil information to the specifications of the end-users in situations where legacy data is not available or representative of the scale to be mapped or when it is prohibitive to conduct new soil survey.

7.4. Future Work

Directly from the investigations carried out during this project there are opportunities for further work; both in terms of filling gaps which this project did not adequately address and more generally for advancing digital soil mapping into the operational phase and beyond.

- There is an increasing richness of data sources describing the physical distribution of the Earth's resources. These data sources retrieved from both remote and proximal sensing systems are being disseminated widely with improved qualities and resolutions. From these data sources, the search is on for *scorpan* factors that may improve SSPFs. The suitability of these data sources for predicting sub-soil distribution of soil properties needs to be investigated.
- Related to the first point are the recent efficiencies gained from spectroscopic techniques for efficiently measuring soil properties. Visible, near-infrared, and mid-infrared diffuse reflectance spectroscopy techniques represent possible alternatives to enhance or replace conventional laboratory methods of soil analysis (Viscarra Rossel et al. 2006). These techniques could be applied to operational DSM in two ways: (1) Coupling these techniques for retrieving previously unrecorded soil property information from existing legacy soil spatial datasets; (2) The efficiencies gained in terms of lower costs and time for measuring soil properties means that the density of samples taken during field soil survey can be increased, meaning more data to calibrate SSPFs.
- The efficacy of the methodologies proposed in this project in delivering relevant spatial soil information can never really be validated until real situations present themselves. One real-world application is Digital Soil Assessment (DSA) (Carre et al. 2007) which is interlinked with DSM where coupling of relevant soil property information with evaluations of soil functions and threats is performed. The mapping of particular soil functions and threats may serve the end-users more adequately than just thematic soil property maps if we are considering issues such as soil protection, and should be investigated further.

Serving the end-user is important and there is a need to specifically address ways in which spatial soil information is distributed and delivered. This is a 'serving' and infrastructure problem that soil mapping agencies will probably only be concerned with. Packaging information and efficient distribution thereof will ensure wider and continued use of spatial soil information.

- This project has concentrated only on use of legacy soil data observations for digital soil mapping. However, using the soil maps created from conventional survey within a digital soil mapping framework will also be invaluable for delivering comprehensive spatial soil information. Conventional soil maps and their legends are representations of the soil surveyor's mental soil-landscape model (Bui 2004). Some sort of knowledge engineering is required to extract this valuable knowledge and couple with building SSPFs (knowledge-based systems). Efficient and continuous methods of conventional soil map disaggregation for retrieving spatial soil property information need also be investigated. Furthermore, statistical data assimilation methods for combining conventional soil map information with digital soil map information (based on soil site observations) could also be investigated. This data assimilation work is practised in climatology research for assimilating data types of different qualities and modes of measurement (Li and Shao 2010), its value for DSM is yet to be realised.
- Improving the predictive power of SSPFs will also improve the certainty of the predicted digital soil maps. Some fine tuning of the uncertainty method as discussed previously would be ideal so as to optimise the quantifications of uncertainty for each target soil property independently.
- Related to the previous point is that further work in necessary for defining perceived thresholds of digital soil map quality in terms of whether a map is suitable for a given purpose. What this requires is collaboration with end-users at all levels—policy makers, researchers, farmers etc— to decide when a digital soil map is acceptable or unacceptable. It is likely such thresholds or criteria will vary somewhat given the purpose and scale of the information required. However, once some criteria have been developed, it is foreseeable there will be strong rationale for motivating further investments in input data,

both in terms of new soil survey and sampling and also exploration or development of new covariate data sources.

- Exploring other methods and sampling designs for the proper validation of digital soil maps need to be investigated. Investigating whether compact geographical randomised sampling is useful for validating a series of thematic maps concurrently within a single mapping domain needs to be investigated.
- A more formal evaluation of the uncertainties of digital soil maps generated from scale manipulations needs to be investigated.
- Optimal methods for validating digital soil maps with block support need to be investigated and tested.

More generally, soil science appears to be back on the research agenda globally (Hartemink and McBratney 2008). The varied functions that soils perform and the human dependence on these functions is driving research to understand soils better and to manage them more appropriately. It is an exciting time to be involved in research that is potentially able to result in positive outcomes both for humans *and* soils. The task now is to disseminate the practicable methodologies proposed in this project to soil mapping agencies in order to meet the global demand for comprehensive spatial soil information.

7.5. References

Addiscott, T.M., 2010. Entropy, non-linearity and hierarchy in ecosystems. Geoderma, 160: 57-63.

Bishop, T.F.A., McBratney, A.B. and Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. Geoderma, 91: 27-45.

Brus, D.J., Spatjens, L. and de Gruijter, J.J., 1999. A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. Geoderma, 89: 129-148.

Bui, E.N., 2004. Soil survey as a knowledge system. Geoderma, 120: 17-26.

Burrough, P.A., 1993. Soil variability: a late 20th century view. Soil and Fertilizers 56: 529-562.

Carre, F., McBratney, A.B., Mayr, T. and Montanarella, L., 2007. Digital soil assessments: Beyond DSM. Geoderma, 142: 69-79.

Crawford, J.W., 2010. Can complex be simple? Geoderma, 160: 1-2.

Cressie, N.A.C., 1993. Statistics for Spatial Data. John Wiley and Sons, New York.

de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P. and Knotters, M., 2006. Sampling for Natural Resource Monitoring. Springer-Verlag, Berlin.

Diggle, P.J., Tawn, J.A. and Moyeed, R.A., 1998. Model-based geostatistics. Journal of the Royal Statistical Society Series C-Applied Statistics, 47: 299-326.

Hartemink, A.E. and McBratney, A., 2008. A soil science renaissance. Geoderma, 148: 123-129.

Heuvelink, G.B.M., 1998. Error Propagation in Environmental Modelling. Taylor and Francis, London.

Kempen, B., Brus, D.J., Heuvelink, G.B.M. and Stoorvogel, J.J., 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. Geoderma, 151: 311-326.

Lark, R.M., Cullis, B.R. and Welham, S.J., 2006. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. European Journal of Soil Science, 57: 787-799.

Li, M. and Shao, Q., 2010. An improved statistical approach to merge satellite rainfall estimates and rain-gauge data. Journal of Hydrology, 385: 51-64.

McBratney, A.B. and de Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science, 43: 159–175.

McKenzie, N.J. and Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma, 89: 67-94.

Minasny, B. and McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosciences, 32: 1378-1388.

Minasny, B., McBratney, A.B. and Lark, R.M., 2008. Digital soil mapping technologies for countries with sparse data infrastructures. In: A.E. Hartemink, A.B. McBratney and M.D. Mendonca-Santos (Editors), Digital Soil Mapping with Limited Data. Springer, Australia, pp. 15-30.

Minasny, B., McBratney, A.B., Malone, B.P. and Wheeler, I., In Review. Digital mapping of soil carbon. Geoderma.

Minasny, B., McBratney, A.B., Mendonca-Santos, M.L., Odeh, I.O.A. and Guyon, B., 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. Australian Journal of Soil Research, 44: 233-244.

Stevens, D.L., Olsen, A.R., 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99: 262-278.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. and Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma, 131: 59-75.

Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Computers & Geosciences 36: 1261-1267.